FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

The Maximum Principle and Controlled Diffusion

MASTERARBEIT

zur Erlangung des akademischen Grades

Master of Science

im Studiengang Mathematik

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

Fakultät für Mathematik und Informatik

eingereicht von Julian Wendt

geb. am 10.08.1993 in Hamburg

Betreuer: Prof. Dr. Stefan Ankirchner

Jena, 4. April 2019

Abstract

The objective of this master thesis is to solve a controlled diffusion problem via Pontryagin's maximum principle. To that end, we review in the first part basic notions that are relevant for the course of this thesis. In the second part, we study a controlled diffusion problem, in which one completely determines the diffusion of the process, but has no direct influence on the drift coefficient. We apply the maximum principle and solve the adjoint forward-backward stochastic differential equation with the help of the method of decoupling fields. The last part presents a connection between the value function and the decoupling fields of our control problem. In particular, we show that the weak derivative of the value function is equal to the decoupling field.

Zusammenfassung

Das Ziel der vorliegenden Masterarbeit ist es, ein kontrolliertes Diffusionsproblem mit dem Pontryaginschen Maximumsprinzip zu lösen. Wir führen daher im ersten Teil grundlegende Begriffe ein, die im Laufe dieser Arbeit von Bedeutung sein werden. Im zweiten Teil lösen wir ein kontrolliertes Diffusionsproblem, in welchem man die Diffusion vollständig bestimmt, man jedoch keinen direkten Einfluss auf den Driftkoeffizienten hat. Dazu wenden wir das Maximumsprinzip an und lösen die adjungierte stochastische Vorwärts-Rückwärts-Differentialgleichung mit der Methode der Entkopplungsfelder. Der letzte Abschnitt befasst sich mit der Verbindung der Wertefunktion mit dem Entkopplungsfeld. Wir zeigen, dass die schwache Ableitung der Wertefunktion gleich dem Entkopplungsfeld ist.

Acknowledgements

I would like to express my gratitude to Professor Dr. Stefan Ankirchner and Dr. Alexander Fromm for supervising my work on this master thesis. Thank you for the guidance and for taking the time to answer all my questions. Also, I thank Verena Köpp for proofreading and reducing the number of errors in this thesis. Finally, I wish to thank my family for supporting me throughout my years of study and through the process of writing this thesis.

Contents

Chapter 1

Introduction

Stochastic control theory has been a very important field in applied mathematics over the past decades. Its ambition is to find an optimal control for a diffusion process such that a cost functional is minimized or maximized. Motivated by various applications in mathematical finance and economics, two main techniques to solve such problems have been developed from the 1950s onwards. The first technique, the classical Hamilton-Jacobi-Bellman (HJB) approach via dynamic programming, characterizes a solution in terms of a partial differential equation (PDE), the so-called HJB equation. The verification theorem implies that every solution to this PDE coincides with the value function if it satisfies certain growth conditions. This theorem additionally provides an optimal control.

The second technique, Pontryagin's maximum principle, states sufficient conditions for the optimality of a control given that there is a solution to the adjoint forward-backward stochastic differential equation (FBSDE). In particular, it converts the task of finding an optimal control into the problem of solving an FBSDE that is in general coupled.

In this thesis we study the control problem presented in [1], but unlike this publication we assume that the control processes are bounded, which means that they take values in a compact interval only. To be more precise, we consider a controlled diffusion process M driven by the stochastic differential equation

$$M_t^\alpha = \mu(t, M_t^\alpha)dt + \alpha_t dW_t,$$

where μ is an affine linear function in M. The controller completely determines the diffusion of the process, but has no direct influence on the drift. Our control problem consists of minimizing the cost functional

$$\mathbb{E}\left[\int_0^T f(s, \alpha_s)\,\mathrm{d}s + g(M_T^\alpha)\right]$$

over a class of suitable controls α. In more detail, the controller tries to minimize the costs by steering the process in the best possible way. Here the functions f and g are especially allowed to depend on the Brownian paths, this means that our framework is non-Markovian what makes the maximum principle for our setting suitable, unlike the classical HJB approach.

In practice such diffusion control problems arise, for example, in portfolio optimization. In this case the process M describes the portfolio value process with volatility α. The function f can then be interpreted as hedging costs that appear by reducing the volatility of the portfolio, while $-g$ can be viewed as a utility function. Another example is presented in the introduction of chapter 3.

As proposed in [1], the above control problem seems to be unsolved in this framework. However, there are recent articles that expand the classical HJB theory to non-Markovian settings considering for example path-dependent PDEs (see e.g. [3] and [4]). The maximum principle avoids such considerations, but confronts us in our control problem with the task of solving a fully coupled FBSDE. In general, there are neither existence nor uniqueness results for global solutions even if certain Lipschitz conditions are fulfilled. But the method of decoupling fields provides a rich theory for our purpose. In particular, we use this technique to show that there exists a solution to the adjoint FBSDE, which enables us to determine an optimal control via the maximum principle.

Introducing the two approaches above leads to the question how they are connected. An extensive study of this question can be found, for instance, in chapter 5 of [13]. In contrast to this book, we link the value function and the so-called decoupling field that is determined by the adjoint FBSDE. To be more precise, we show that the weak derivative of the value function is equal to the decoupling field.

This thesis is organized as follows: In chapter 2 we review basic notions about BSDEs and FBSDEs. We summarize the main statements of the method of decoupling fields and about weak derivatives. Moreover, we present Pontryagin's maximum principle for a non-Markovian framework. In chapter 3 we study our main topic of this thesis. We prove that the control problem aforementioned has a solution using the method of decoupling fields. Finally, in chapter 4 we connect the HJB approach with the FBSDE approach of the maximum principle. We show how the value function and the decoupling field are linked in the general non-Markovian setting.

Chapter 2

Preliminaries

Pontryagin and his team introduced in the 1950s the so-called *maximum principle* to solve optimal control problems. They first studied deterministic problems, but later on also stochastic control problems were considered, starting with works of Bismut in 1973. He studied uncoupled FBSDEs with a linear backward equation, which led to the research on *backward stochastic differential equations* and *forward-backward stochastic differential equations*, abbreviated by BSDEs and FBSDEs, respectively (cf. [13, p. 101ff]).

In the following decades much research on BSDEs and FBSDEs in the context with the maximum principle and on the connection to PDEs was done. However, there were no general results about the existence and uniqueness of solutions until Pardoux and Peng introduced in 1990 the theory of general BSDEs. Later on in the early 90s, studies on coupled FBSDEs began (cf. [5, p. 17] and [9, p. vii]). Nowadays, BSDEs and FBSDE are an important field in stochastics, for example, because of their application in partial differential equations, mathematical finance and stochastic control theory (see e.g. [11]).

In this chapter we introduce the aforementioned terms BSDE and FBSDE. We present basic facts about existence and uniqueness of solutions and connected with that we demonstrate the *method of decoupling fields*, which is very useful in constructing solutions to FBSDEs. As it turns out, Lipschitz continuity of the generator of a BSDE ensures the existence and uniqueness of a solution. In the case of FBSDEs, however, Lipschitz conditions are not sufficient to achieve even solvability. At this point the method of decoupling fields allows us to construct solutions on sufficiently small time intervals under certain *standard Lipschitz conditions*. This technique uses extensively the notion of weak differentiability and we therefore introduce it in section 2.5. We state and prove results for these derivatives that are relevant for this thesis.

In section 2.4 we explain the concept of *Pontryagin's maximum principle*, which plays an important role in stochastic control because it offers a technique for determining optimal controls. Moreover, it connects the optimal control theory with the theory of FBSDEs, because this approach transforms the task of finding an optimal control into the task of solving a possibly coupled FBSDE depending on a control. If there exists a solution and the same control process minimizes the so-called *Hamiltonian*, the maximum principle states that this control is optimal.

Finally, we emphasize that all the theory presented in this chapter can be generalized to the multidimensional case. The statements stay basically the same and just some small adjustments have to be made. For instance, one has to consider vector and matrix norms instead of absolute values.

Throughout this chapter we assume the following: Let $T > 0$ be a deterministic time horizon and $W = (W_t)_{t\in[0,T]}$ be a Brownian motion on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, where $(\mathcal{F}_t)_{t\in[0,T]}$ is the natural filtration of W. In more detail, the filtration is defined by $\mathcal{F}_t := \sigma\left(\mathcal{N}, (W_s)_{s\in[0,t]}\right)$ with $\mathcal{N}$ denoting the set of all $\mathbb{P}$-null sets. Additionally, we define for $t \in [0,T]$ the process spaces

$$H^2_{t,T} := \left\{ X : \Omega \times [t,T] \to \mathbb{R} \;\middle|\; X \text{ is progr. mb. and } \mathbb{E}\left[\int_t^T X_s^2 \,\mathrm{d}s\right] < \infty \right\},$$
$$S^2_{t,T} := \left\{ X : \Omega \times [t,T] \to \mathbb{R} \;\middle|\; X \text{ is progr. mb. and } \mathbb{E}\left[\sup_{s\in[t,T]} X_s^2\right] < \infty \right\},$$

and we denote by $\mathcal{H}^2_{t,T}$ and $\mathcal{S}^2_{t,T}$ the corresponding quotient spaces with respect to the equivalence relations $\sim_H$ and $\sim_S$, respectively, given by

$$X \sim_H Y :\Longleftrightarrow \mathbb{P} \otimes \lambda\left(\{(\omega,t) \in \Omega \times [0,T] : X_t(\omega) \neq Y_t(\omega)\}\right) = 0,$$
$$X \sim_S Y :\Longleftrightarrow X \text{ and } Y \text{ are indistinguishable.}$$

2.1 Backward stochastic differential equations

In this section we introduce *backward stochastic differential equations* (BSDEs) and present some important facts. Throughout we are going to follow chapter 6.2.1 of [11].

Let $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a measurable function that is progressively measurable, meaning that for all $t \in [0,T]$ the mapping f restricted to $\Omega \times [0,t] \times \mathbb{R} \times \mathbb{R}$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t]) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$-measurable. Furthermore, let ξ be an $\mathcal{F}_T$-measurable random variable. A BSDE on $[0,T]$ is an equation of the form

$$-dY_t = f(t, Y_t, Z_t)dt - Z_t dW_t, \; t \in [0,T], \quad Y_T = \xi,$$

or equivalently,

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\,\mathrm{d}s - \int_t^T Z_s \,\mathrm{d}W_s, \; t \in [0,T]. \tag{2.1}$$

We call the function f the *generator* and the random variable ξ the *terminal condition*. We refer to the pair (ξ, f) as the *parameters* of the BSDE (2.1). We call the parameters (ξ, f) *standard*, if

- ξ has finite second moment, i.e. $\mathbb{E}\xi^2 < \infty$,
- $f(\cdot,\cdot,0,0) \in \mathcal{H}^2_{0,T}$,
- f is Lipschitz continuous in (y,z) in the sense that there exists a constant $L \geq 0$ such that for all $y_1, y_2, z_1, z_2 \in \mathbb{R}$ it holds that

  $$|f(\omega,t,y_1,z_1) - f(\omega,t,y_2,z_2)| \leq L\left(|y_1 - y_2| + |z_1 - z_2|\right)$$

  for $\mathbb{P} \otimes \lambda$-almost all $(\omega,t) \in \Omega \times [0,T]$.

Definition 2.1. We say that $(Y, Z) \in \mathcal{S}^2_{0,T} \times \mathcal{H}^2_{0,T}$ is a solution to the BSDE (2.1) if

- $\int_0^T |f(s, Y_s, Z_s)| \,\mathrm{d}s < \infty$, $\mathbb{P}$-almost surely, and
- equation (2.1) is satisfied for all $t \in [0, T]$, $\mathbb{P}$-almost surely.

We say that a solution is unique if for two solutions (Y, Z) and $(\tilde{Y}, \tilde{Z})$, the processes Y and $\tilde{Y}$ are indistinguishable, and the processes Z and $\tilde{Z}$ coincide $\mathbb{P} \otimes \lambda$-almost everywhere.

Remark 2.2. Note that solutions to BSDEs are only uniquely defined in the following sense: Given a solution (Y, Z) to (2.1) in the space $\mathcal{S}^2_{0,T} \times \mathcal{H}^2_{0,T}$, one has that also $(\tilde{Y}, \tilde{Z})$ solves the BSDE (2.1) if the processes Y and $\tilde{Y}$ are indistinguishable, and the processes Z and $\tilde{Z}$ are $\mathbb{P} \otimes \lambda$-almost everywhere equal.

For a solution $(Y, Z) \in \mathcal{S}^2_{0,T} \times \mathcal{H}^2_{0,T}$ we can always assume that the process Y is continuous in time due to the form of (2.1). It is indeed enough to redefine Y as the right hand side of (2.1), which yields an indistinguishable version of Y that is continuous.

Theorem 2.3. *Let (ξ, f) be standard parameters. Then there exists a unique solution $(Y, Z) \in \mathcal{S}^2_{0,T} \times \mathcal{H}^2_{0,T}$ to the BSDE (2.1).*

Sketch of the proof. We summarize the proof of this statement presented in Theorem 6.2.1 of [11], which is based on a fixed point method. We define the mapping $\Phi : \mathcal{S}^2_{0,T} \times \mathcal{H}^2_{0,T} \to \mathcal{S}^2_{0,T} \times \mathcal{H}^2_{0,T}$, $(U, V) \mapsto (Y, Z)$, where (Y, Z) is a solution of

$$Y_t = \xi + \int_t^T f(s, U_s, V_s) \,\mathrm{d}s - \int_t^T Z_s \,\mathrm{d}W_s,\; t \in [0, T]. \tag{2.2}$$

We construct (Y, Z) in the followings steps:

- Define the martingale M given by
$$M_t := \mathbb{E}\left[\xi + \int_0^T f(s, U_S, V_s) \,\mathrm{d}s \middle| \mathcal{F}_t\right],\; t \in [0, T].$$
Note that M is square-integrable by our assumptions on (ξ, f), i.e. $\mathbb{E} \int_0^T M_t^2 \,\mathrm{d}t < \infty$.
- Applying the martingale representation theorem to M yields the existence and uniqueness of a process $Z \in \mathcal{H}^2_{0,T}$ with
$$M_t = M_0 + \int_0^t Z_s \,\mathrm{d}W_s,$$
see e.g. Theorem 1.2.9 in [11]. Note that Z is uniquely determined in the sense that for another process $\widetilde{Z}$ with the above property we have
$$\mathbb{P} \otimes \lambda \left(\left\{ (\omega, t) \in \Omega \times [0, T] : Z_t(\omega) \neq \widetilde{Z}_t(\omega) \right\} \right) = 0.$$
- Define the process Y by setting
$$Y_t := \mathbb{E}\left[\xi + \int_t^T f(s, U_s, V_s) \,\mathrm{d}s \middle| \mathcal{F}_t\right] = M_t - \int_0^t f(s, U_s, V_s) \,\mathrm{d}s,\; t \in [0, T].$$
The last equality holds since $(\omega, s) \mapsto f(s, U_s, V_s)$ is progressively measurable, which implies, in particular, that $\int_0^t f(s, U_s, V_s) \,\mathrm{d}s$ is $\mathcal{F}_t$-measurable.

Note that $Y_T = \xi$. By Doob's inequality one can show that $Y \in \mathcal{S}^2_{0,T}$. Furthermore, Y is uniquely defined up to indistinguishability and we can write M as

$$M_t = M_T - \int_t^T Z_s \,\mathrm{d}W_s = \xi + \int_0^T f(s, U_s, V_s)\,\mathrm{d}s - \int_t^T Z_s \,\mathrm{d}W_s, \; t \in [0,T].$$

Consequently, Y satisfies (2.2) and the mapping Φ is well-defined. We observe from the construction above that a solution to BSDE (2.1) is a fixed point of the function Φ, and the other way around. Now one can show that Φ is a contraction on the Banach space $\mathcal{S}^2_{0,T} \times \mathcal{H}^2_{0,T}$ with norm

$$\|(Y,Z)\| = \left(\mathbb{E}\left[\int_0^T e^{\beta s}\left(Y_s^2 + Z_s^2\right)\,\mathrm{d}s\right]\right)^{\frac{1}{2}}, \; (Y,Z) \in \mathcal{S}^2_{0,T} \times \mathcal{H}^2_{0,T},$$

for a certain fixed $\beta > 0$ depending on f. Hence Banach's fixed point theorem implies the existence of a unique fixed point, which is the solution to the BSDE (2.1). □

Remark 2.4. We emphasize the fact that one can write BSDE (2.1) as a forward equation:

$$Y_t = Y_0 - \int_0^t f(s, Y_s, Z_s)\,\mathrm{d}s + \int_0^t Z_s \,\mathrm{d}W_s.$$

This allows us to apply the product formula.

The next and final statement admits a comparison of solutions to BSDEs. In applications it sometimes enables us to show that a solution of a BSDE is bounded if we choose the comparing solution in the right way (see e.g. the proof of Lemma 3.14).

Theorem 2.5 (Comparison principle)**.** *Let (ξ^1, f^1) and (ξ^2, f^2) be standard parameters and let (Y^1, Z^1) and (Y^2, Z^2) be solutions to the corresponding BSDEs. Furthermore, assume that:*

- $\xi^1 \le \xi^2$ *a.s.,*
- $f^1(t, Y_t^1, Z_t^1) \le f^2(t, Y_t^1, Z_t^1)$ *for $\mathbb{P} \otimes \lambda$-almost every $(\omega, t) \in \Omega \times [0,T]$,*
- $f^2(\cdot, Y_\cdot^1, Z_\cdot^1) \in \mathcal{H}^2_{0,T}$.

Then $Y_t^1 \le Y_t^2$ for all $t \in [0,T]$, $\mathbb{P}$-almost surely.

A proof of this theorem can be found in Theorem 6.2.2 of [11].

2.2 Forward-backward stochastic differential equations

This section introduces the notion of *forward-backward stochastic differential equations* (FBSDEs). We define a solution of an FBSDE and present an example of a non-solvable FBSDE. The following is based on [6] and [9].

First of all, let $\mu, \sigma, f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be measurable functions that are progressively measurable, meaning that for all $t \in [0,T]$ the functions μ, σ, f restricted to $\Omega \times [0,t] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ are $\mathcal{F}_t \otimes \mathcal{B}([0,t]) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$-measurable. Additionally, let $\xi : \Omega \times \mathbb{R} \to \mathbb{R}$ be an $\mathcal{F}_T \otimes \mathcal{B}(\mathbb{R})$-measurable function and let $t \in [0,T]$. An FBSDE on the interval $[t,T]$ is a system

$$\begin{aligned} X_s &= x + \int_t^s \mu(r, X_r, Y_r, Z_r)\,\mathrm{d}r + \int_t^s \sigma(r, X_r, Y_r, Z_r)\,\mathrm{d}W_r, \\ Y_s &= \xi(X_T) - \int_s^T f(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_s^T Z_r\,\mathrm{d}W_r, \ \ s \in [t,T], \end{aligned} \tag{2.3}$$

consisting of a forward equation starting in $x \in \mathbb{R}$ at time t and a backward equation with terminal condition ξ. We call (ξ, μ, σ, f) the *parameters* of the FBSDE (2.3). If the parameters (ξ, μ, σ, f) do not depend on ω, we call the FBSDE (2.3) *Markovian.* Otherwise we call it *non-Markovian.* In this thesis we consider primarily the non-Markovian case. Nevertheless, we define in either cases a solution as follows.

Definition 2.6. We say that a triple $(X, Y, Z) \in \mathcal{S}^2_{t,T} \times \mathcal{S}^2_{t,T} \times \mathcal{H}^2_{t,T}$ is a solution to the FBSDE (2.3) on $[t,T]$ with parameters (ξ, μ, σ, f) and initial value $x \in \mathbb{R}$ if

(1) $\int_t^T |\mu(s, X_s, Y_s, Z_s)|\,\mathrm{d}s, \int_t^T |f(s, X_s, Y_s, Z_s)|\,\mathrm{d}s, \int_t^T |\sigma(s, X_s, Y_s, Z_s)|^2\,\mathrm{d}s < \infty$, a.s.,

(2) equation (2.3) is satisfied for all $s \in [t,T]$, $\mathbb{P}$-almost surely.

We say that a solution is unique if for two solutions $(X, Y, Z), (\tilde{X}, \tilde{Y}, \tilde{Z})$, the processes $X, \tilde{X}$ and $Y, \tilde{Y}$ are indistinguishable, respectively, and $Z, \tilde{Z}$ coincide $\mathbb{P} \otimes \lambda$-almost everywhere.

Remark 2.7. The processes X and Y can be assumed to be continuous, because by redefining these processes as the right hand sides of (2.3), respectively, one obtains continuous versions of X and Y.

We emphasize that it is possible to define a solution to an FBSDE in a weaker sense, for instance, one can just require that $X, Y, Z \in \mathcal{H}^2_{t,T}$, that (1) holds true and that (2.3) is satisfied almost surely for all $s \in [t,T]$. In section 2.3 we will work with this weaker definition of a solution, because the method of decoupling fields provides solutions in this sense only. However, we will show that there is also a solution in the sense of Definition 2.6.

For studying the solvabilty of an FBSDE it is useful two distinguish *coupled* and *decoupled* systems. The latter means that either the forward or the backward equation does not depend on the other one. In other words, the FBSDE (2.3) is said to be *decoupled* if either μ and σ do not depend on (y, z), or f and ξ do not depend on x. Otherwise it is called *coupled*. In the decoupled setting one can try to solve the FBSDE by first solving the independent equation, and by plugging this solution into the remaining one. In this case one can try to apply the existence and uniqueness results for SDEs and BSDEs.

Solving coupled FBSDEs, however, turns out to be much harder. The usual assumptions on the parameters of the equation, for instance, Lipschitz continuity, do not ensure existence and uniqueness of a solution, in contrast to the case of SDEs and BSDEs. In fact, it can happen that an FBSDE has no solution at all like presented in the next example, which is discussed in Example 2.3.2 of [6] and in Proposition 3.1 of [9] in a similar way.

Example 2.8. Let $T = 1$ and $t \in [0, T)$. Consider for $x \in \mathbb{R}$ the coupled FBSDE on $[t, T]$

$$\begin{aligned} X_s &= x + \int_t^s Y_r \,\mathrm{d}r, \\ Y_s &= X_T - \int_s^T Z_r \,\mathrm{d}W_r, \;\; s \in [t, T]. \end{aligned} \tag{2.4}$$

We see that here $\mu(y) = y, \xi(y) = y$ and $\sigma, f = 0$. All of these functions are obviously Lipschitz continuous and have only linear growth, but there does not have to exist a solution to (2.4) as we demonstrate below. Here the crucial point is the choice of t. First of all, note that for any $t \in [0, T]$ the processes $X, Y, Z \equiv 0$ solve the above FBSDE on $[t, T]$ if $x = 0$. Therefore, we assume in the following that $x \neq 0$.

In the case of $t = 0$, there does not exist a solution as one can show as follows. Assume on the contrary that there exists a solution (X, Y, Z) to (2.4) on $[0, T]$. This means that we have $X, Y \in \mathcal{S}^2_{0,T}, Z \in \mathcal{H}^2_{0,T}$ and therefore $\mathbb{E} \int_s^T Z_r \,\mathrm{d}W_r = 0$ for all $s \in [0, T]$. In addition, we can apply dominated convergence to show that the functions $\hat{x}$ and $\hat{y}$, given by $\hat{x}(s) := \mathbb{E}X_s,\ \hat{y}(s) := \mathbb{E}Y_s,\ s \in [0, T]$, satisfy the ordinary differential equation

$$\hat{x}'(s) = \hat{y}(s), \qquad \hat{y}'(s) = 0, \qquad s \in [0, T], \tag{2.5}$$

with the boundary conditions $\hat{x}(0) = x$ and $\hat{y}(T) = \hat{x}(T)$. But a general solution $(\tilde{x}, \tilde{y})$ to the ordinary differential equation (2.5) with terminal condition $\tilde{x}(T) = \tilde{y}(T)$ has the form

$$\tilde{x}(s) = cs, \quad \tilde{y}(s) = c, \quad s \in [0, T]$$

for some constant $c \in \mathbb{R}$. Moreover, we observe that $\tilde{x}(0) = 0 \neq x$. Thus, there cannot exist a solution to (2.5) with these boundary conditions and hence our assumption has to be wrong. Consequently, the FBSDE does not have a solution on $[0, T]$. Another possible way of proving the above statement is with the help of decoupling fields. In Example 2.23 we present this in more detail.

In the case of $t \in (0, T]$ we actually find a solution to the FBSDE (2.4) on $[t, T] \subsetneq [0, T]$ by observing that for $Z \equiv 0$ the FBSDE is just an ordinary differential equation. Its solution can be calculated and thus we obtain by using the boundary conditions

$$X_s = x\frac{s}{t}, \quad Y_s = \frac{x}{t}, \quad Z_s = 0, \; s \in [t, T].$$

Conversely, one can verify that (X, Y, Z), defined in this way, solves indeed the FBSDE (2.4) on $[t, T]$.

We emphasize that this example does not rely on the particular choice of T. Even assuming that $T \in [1, \infty)$ suffices to show, in the same manner as above, that the FBSDE (2.4) is for $t \in (T - 1, T]$ solvable and for $t = T - 1$ non-solvable.

As one can observe in the example above, assumptions on the parameters are not sufficient to gain solvability of an FBSDE. Structural assumptions have to be made as well, since in some cases one can get solutions if T is small enough depending on the parameters (see e.g. Theorem 5.1 in [9]). There do exist techniques to solve coupled FBSDEs, for instance, the *Four Step Scheme*, the *Method of Continuation* and the *Contraction Method*, to name a few. For each of these approaches one has to make significant restrictions to the class of possible parameter functions. In the Four Step Scheme, for instance, only the Markovian case with sufficiently smooth parameters is considered. For more details on these methods we refer to the literature (e.g. [9], [13]).

Another approach to study coupled FBSDEs is the *method of decoupling fields*, on which we focus in this thesis. We introduce this method in the next section. Unlike the aforementioned concepts, this technique is also concerned with the existence of solutions on smaller time intervals, i.e. intervals of the form $[t,T] \subseteq [0,T]$. Additionally, it somehow treats the more general case in the sense that it only requires the so-called *standard Lipschitz conditions*, described below in section 2.3.

2.3 Method of decoupling fields

We present in this section the *method of decoupling fields* developed in [6]. Therefore, the following is based on that thesis. We emphasize that all definitions, statements and proofs can be found there.

The method of decoupling fields deals with the solvability of coupled FBSDEs, which can be non-Markovian. This approach does not only search for global solutions of FBSDEs described in Definition 2.6, but is also interested in solutions on smaller time intervals and especially in finding a random function u of time and space, called the *decoupling field*, that connects the forward and backward equation in the sense that $Y_t = u(t, X_t)$. To find such a function, one divides the interval $[0,T]$ into finitely many small intervals and then tries to construct decoupling fields on each of them by going from the right boundary to the left. In this procedure one only has to require that X, Y, Z exist locally on small intervals instead of requiring global existence on $[0,T]$. The crucial point in this construction is that one can concatenate decoupling fields to obtain a decoupling field on a larger interval, and that one can receive the processes X, Y, Z at the same time. To apply this technique one has to require the so-called *standard Lipschitz conditions* of the parameters.

Consider now the setting of section 2.2, this means that we define the measurable functions $\mu, \sigma, f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ that are progressively measurable, meaning that their restriction to $\Omega \times [0,t] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t]) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$-measurable for all $t \in [0,T]$. Additionally, let $\xi : \Omega \times \mathbb{R} \to \mathbb{R}$ be an $\mathcal{F}_T \otimes \mathcal{B}(\mathbb{R})$-measurable function and $x \in \mathbb{R}$. We consider the FBSDE

$$\begin{aligned} X_t &= x + \int_0^t \mu(s, X_s, Y_s, Z_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s, Y_s, Z_s)\,\mathrm{d}W_s, \\ Y_t &= \xi(X_T) - \int_t^T f(s, X_s, Y_s, Z_s)\,\mathrm{d}s - \int_t^T Z_s\,\mathrm{d}W_s, \end{aligned} \tag{2.6}$$

for $t \in [0,T]$. In this section we always make the above assumptions without explicitly mentioning them. Furthermore, we emphasize that we often call a triple (X, Y, Z) a solution to an FBSDE even if it just satisfies the weaker definition aforementioned in Remark 2.7. The reason for that is the construction of a solution in the proof of Theorem 2.16, which provides only a solution in this sense. However, in the final statement (Theorem 2.21) we show that there is also a solution to (2.6) in the sense of Definition 2.6.

Definition 2.9. Let $t \in [0,T]$ and $u : \Omega \times [t,T] \times \mathbb{R} \to \mathbb{R}$. We say that the function u is a *decoupling field* for (ξ, μ, σ, f) on $[t,T]$ if

(1) $u(T,\cdot) = \xi(\cdot)$ a.s.,

and for all $t_1, t_2 \in [t,T], t_1 \leq t_2$, and any $\mathcal{F}_{t_1}$-measurable $X_{t_1} : \Omega \to \mathbb{R}$ there exist progressively measurable processes X, Y, Z on $[t_1, t_2]$ such that for all $s \in [t_1, t_2]$

(2) $\displaystyle\int_{t_1}^{t_2} |\mu(r, X_r, Y_r, Z_r)| \,\mathrm{d}r, \int_{t_1}^{t_2} |f(r, X_r, Y_r, Z_r)| \,\mathrm{d}r, \int_{t_1}^{t_2} |\sigma(r, X_r, Y_r, Z_r)|^2 \,\mathrm{d}s < \infty$ a.s.,

(3) $\displaystyle X_s = X_{t_1} + \int_{t_1}^{s} \mu(r, X_r, Y_r, Z_r) \,\mathrm{d}r + \int_{t_1}^{s} \sigma(r, X_r, Y_r, Z_r) \,\mathrm{d}W_r$ a.s.,

(4) $\displaystyle Y_s = Y_{t_2} - \int_{s}^{t_2} f(r, X_r, Y_r, Z_r) \,\mathrm{d}r - \int_{s}^{t_2} Z_r \,\mathrm{d}W_r$ a.s.,

(5) $Y_s = u(s, X_s)$ a.s.

Remark 2.10. Concerning the above definition we want to point out the following:

- We refer to (5) as the *decoupling condition*.
- The properties (2)-(4) mean that (X, Y, Z) solves the FBSDE (2.6) on $[t_1, t_2]$ with initial condition X_{t_1}.
- Since all the equalities above are almost surely true, a modification of a decoupling field is again a decoupling field to the same problem. Furthermore, considering that we always require progressive measurability for solutions to FBSDEs leads to the question whether there exists a progressively measurable modification of the decoupling field. The answer is indeed positive if the decoupling field is Lipschitz continuous for almost all $\omega \in \Omega$. See Remark 2.14 for details.
- By using X_{t_1} to denote the initial value we make a slight abuse of notation. However, since X fulfils the FBSDE almost surely, we know that the process X coincides with this value almost surely at time t_1 and therefore this definition makes sense.
- The processes X, Y, Z arising in Definition 2.9 are not required to be unique in some sense for given t_1, t_2 and X_{t_1}.

As already mentioned, the concatenation of two decoupling fields is again a decoupling field. We present this useful property in the following lemma.

Lemma 2.11. *Let $s, t \in [0,T), s < t$. If u_2 is a decoupling field for (ξ, μ, σ, f) on $[t,T]$ and u_1 is a decoupling field for $(u_2(t,\cdot), \mu, \sigma, f)$ on $[s,t]$, then the function u, defined by*

$$u(\omega, r, x) := \begin{cases} u_1(\omega, r, x) & , r \in [s,t] \\ u_2(\omega, r, x) & , r \in (t,T] \end{cases},$$

for $(\omega, r, x) \in \Omega \times [s,T] \times \mathbb{R}$, is a decoupling field for (ξ, μ, σ, f) on $[s,T]$.

The complete proof of this lemma can be found in Lemma 2.1.2 of [6]. We just sketch the main steps.

Sketch of the proof. We have to check whether u satisfies Definition 2.9.

Let $t_1, t_2 \in [s,T]$ such that $t_1 \leq t_2$. If either $t_1 \geq t$ or $t_2 \leq t$, the required property follows from u_1 and u_2 being decoupling fields. Consequently, we assume that $t_1 \in [s,t)$ and $t_2 \in (t,T]$. Let, moreover, $\tilde{X}_{t_1} : \Omega \to \mathbb{R}$ be $\mathcal{F}_{t_1}$-measurable. Now according to the definition of the decoupling fields u_1 and u_2 the following holds true:

1. There exists a solution $(\tilde{X}, \tilde{Y}, \tilde{Z})$ to the FBSDE on $[t_1, t]$ with initial condition $\tilde{X}_{t_1}$, terminal condition $u_2(t, \tilde{X}_t)$ and decoupling condition $\tilde{Y}_r = u_1(r, \tilde{X}_r) = u(r, \tilde{X}_r)$ for all $r \in [t_1, t]$.

2. There exists a solution $(\hat{X}, \hat{Y}, \hat{Z})$ to the FBSDE on $[t, t_2]$ with initial condition $\tilde{X}_t$, terminal condition $\xi(\hat{X}_{t_2})$ and decoupling condition $\hat{Y}_r = u_2(r, \hat{X}_r)$ for all $r \in [t, t_2]$.

This construction yields $\tilde{X}_t = \hat{X}_t$ and $\tilde{Y}_t = u_1(t, \tilde{X}_t) = u_2(t, \hat{X}_t) = \hat{Y}_t$ a.s. If we define the processes X, Y, Z by

$$X := \tilde{X}\mathbb{1}_{[t_1,t]} + \hat{X}\mathbb{1}_{(t,t_2]}, \quad Y := \tilde{Y}\mathbb{1}_{[t_1,t]} + \hat{Y}\mathbb{1}_{(t,t_2]} \quad \text{and} \quad Z := \tilde{Z}\mathbb{1}_{[t_1,t]} + \hat{Z}\mathbb{1}_{(t,t_2]},$$

it is straightforward to verify that the decoupling condition is satisfied and that (X, Y, Z) solves the FBSDE on $[t_1, t_2]$. □

We now introduce the following terms and abbreviations:

- We denote by $L_{\sigma,z}$ the Lipschitz constant of σ w.r.t. the last component, i.e.

$$L_{\sigma,z} := \sup_{t\in[0,T]} \inf\{L \geq 0 \mid \text{for almost all } \omega \in \Omega : \forall x, y, z_1, z_2 \in \mathbb{R}$$
$$|\sigma(t,x,y,z_1) - \sigma(t,x,y,z_2)| \leq L|z_1 - z_2|\},$$

  and $L_{\sigma,z}^{-1}$ denotes

$$L_{\sigma,z}^{-1} = \begin{cases} \frac{1}{L_{\sigma,z}} & , L_{\sigma,z} \in (0,\infty) \\ \infty & , \text{else.} \end{cases}$$

  In a similar way we define for a function $u : \Omega \times I \times \mathbb{R} \to \mathbb{R}$ the Lipschitz constant of u in x by $L_{u,x}$, i.e.

$$L_{u,x} := \sup_{t\in I} \inf\{L \geq 0 \mid \forall x_1, x_2 \in \mathbb{R} : |u(t,x_1) - u(t,x_2)| \leq L|x_1 - x_2|, \text{ a.s.}\},$$

  where $I \subseteq [0,T]$ is an interval.

- For $t \in [0,T]$ and an integrable random variable X we define $\mathbb{E}_t[X] := \mathbb{E}[X \mid \mathcal{F}_t]$ and

$$\mathbb{E}_{t,\infty}[X] := \operatorname{ess\,sup} \mathbb{E}[X \mid \mathcal{F}_t] = \inf\{c \in [-\infty,\infty] : \mathbb{E}[X \mid \mathcal{F}_t] \leq c \text{ a.s.}\}.$$

- For a measurable function f we denote by $\|f\|_\infty$ the essential supremum of $|f|$.

- The symbol ∂_x denotes weak differentiation w.r.t. x. For details on this subject consult section 2.5.

To restrict the class of possible parameter functions for FBSDEs, we introduce the so-called *standard Lipschitz conditions*. Under these requirements we can develop a theory for decoupling fields and therefore also for solutions to FBSDEs in the general non-Markovian case, as we will see later on.

Definition 2.12. We say that the parameters (ξ, μ, σ, f) satisfy *standard Lipschitz conditions*, abbreviated by SLC, if

(1) μ, σ, f are Lipschitz continuous in (x, y, z) with Lipschitz constant L, i.e. for all $t \in [0, T]$ and almost all $\omega \in \Omega$ we have

$$\begin{aligned}
|\mu(t, x_1, y_1, z_1) - \mu(t, x_2, y_2, z_2)| &\leq L(|x_1 - x_2| + |y_1 - y_2| + |z_1 - z_2|),\\
|\sigma(t, x_1, y_1, z_1) - \sigma(t, x_2, y_2, z_2)| &\leq L(|x_1 - x_2| + |y_1 - y_2| + |z_1 - z_2|),\\
|f(t, x_1, y_1, z_1) - f(t, x_2, y_2, z_2)| &\leq L(|x_1 - x_2| + |y_1 - y_2| + |z_1 - z_2|),
\end{aligned}$$

for all $x_1, x_2, y_1, y_2, z_1, z_2 \in \mathbb{R}$.

(2) $\|\mu(\cdot, \cdot, 0, 0, 0)\|_\infty + \|\sigma(\cdot, \cdot, 0, 0, 0)\|_\infty + \|f(\cdot, \cdot, 0, 0, 0)\|_\infty < \infty$,

(3) $\|\xi(\cdot, 0)\|_\infty < \infty$,

(4) $L_{\xi,x} < L_{\sigma,z}^{-1}$.

Note that the property (4) implies that ξ is Lipschitz continuous in x for almost all $\omega \in \Omega$. The next definitions introduce two notions of regularity for decoupling fields.

Definition 2.13 (weak regularity)**.** For $t \in [0, T)$ a decoupling field $u : \Omega \times [t, T] \times \mathbb{R} \to \mathbb{R}$ for (ξ, μ, σ, f) is called *weakly regular*, if $L_{u,x} < L_{\sigma,z}^{-1}$ and $\sup_{s \in [t,T]} \|u(\cdot, s, 0)\|_\infty < \infty$.

Remark 2.14. A weakly regular decoupling field u is weakly differentiable w.r.t. x, i.e. for almost all $\omega \in \Omega$ and all $s \in [t, T]$ the mapping $u(\omega, s, \cdot)$ is weakly differentiable. Moreover, u can be assumed to be progressively measurable and Lipschitz continuous for all $\omega \in \Omega$, because there exists a modification $\tilde{u}$ having these properties. This modification even satisfies that the function $x \mapsto \tilde{u}(\omega, s, x)$ is weakly differentiable w.r.t. x for all fixed $(\omega, s) \in \Omega \times [t, T]$. With modification we mean that for all $s \in [t, T]$ it holds $u(\omega, s, \cdot) = \tilde{u}(\omega, s, \cdot)$ for almost all $\omega \in \Omega$. These properties are consequences of Lemma 2.1.3 and Lemma 2.1.4 in [6].

To gain more information about the processes X, Y, Z appearing in the FBSDE (2.6), more requirements are necessary.

Definition 2.15 (strong regularity)**.** Let $t \in [0, T)$. A weakly regular decoupling field $u : \Omega \times [t, T] \times \mathbb{R} \to \mathbb{R}$ for (ξ, μ, σ, f) is called *strongly regular* if for all $t_1, t_2 \in [t, T]$, $t_1 \leq t_2$, the processes X, Y, Z arising in Definition 2.9 are $\mathbb{P} \otimes \lambda$-almost everywhere uniquely determined for each *constant* initial value $X_{t_1} = x \in \mathbb{R}$, and satisfy:

(1) For all $x \in \mathbb{R}$

$$\sup_{s \in [t_1, t_2]} \mathbb{E}_{t_1,\infty}\left[X_s^2\right] + \sup_{s \in [t_1, t_2]} \mathbb{E}_{t_1,\infty}\left[Y_s^2\right] + \mathbb{E}_{t_1,\infty}\left[\int_{t_1}^{t_2} Z_s^2 \,\mathrm{d}s\right] < \infty.$$

(2) The processes X, Y, Z with additional dependence on the initial value are measurable. To be more precise, X, Y, Z viewed as mappings $\Omega \times [t,T] \times \mathbb{R} \to \mathbb{R}$ are $\mathcal{F} \otimes \mathcal{B}([t,T]) \otimes \mathcal{B}(\mathbb{R})$-measurable. Furthermore, the processes X, Y, Z are weakly differentiable w.r.t. x such that the random variables X_s and Y_s are weakly differentiable w.r.t. x for every $s \in [t_1, t_2]$. Additionally, we have for the weak derivatives

$$\begin{aligned}
&\operatorname*{ess\,sup}_{x\in\mathbb{R}} \sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}\left[(\partial_x X_s)^2\right] < \infty, \\
&\operatorname*{ess\,sup}_{x\in\mathbb{R}} \sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}\left[(\partial_x Y_s)^2\right] < \infty, \\
&\operatorname*{ess\,sup}_{x\in\mathbb{R}} \mathbb{E}_{t_1,\infty}\left[\int_{t_1}^{t_2} (\partial_x Z_s)^2 \,\mathrm{d}s\right] < \infty.
\end{aligned} \tag{2.7}$$

Regarding (2) we want to point out that we weakly differentiate X, Y, Z w.r.t. the initial value x of the forward equation. The processes Y and Z usually also depend on x due to the coupled nature of the FBSDE. However, we often just write X, Y, Z instead of X^x, Y^x, Z^x.

An important statement for developing the method of decoupling fields is the next theorem (cf. Theorem 2.2.1 in [6]). It presents a technique for constructing decoupling fields for a given FBSDE. However, this construction only works if one considers the FBSDE on a sufficiently small time interval $[t,T]$.

Theorem 2.16. *Let the parameters (ξ, μ, σ, f) satisfy SLC. Then there exists a time $t \in [0,T)$ such that (ξ, μ, σ, f) has a unique (up to modifications) decoupling field u on $[t,T]$ that is weakly regular.*

Sketch of the proof. The detailed proof of this theorem can be found in [6, pp. 20-37]. Due to its complexity we just present the main steps.

1. Let for some $t \in [0,T)$ the mapping $X_t : \Omega \times \mathbb{R} \to \mathbb{R}$ be $\mathcal{F}_t \otimes \mathcal{B}(\mathbb{R})$-measurable such that for a fixed $\hat{t} \in [0,t]$

$$\operatorname*{ess\,sup}_{x\in\mathbb{R}} \mathbb{E}_{\hat{t},\infty}\left[(\partial_x X_t(\cdot,x))^2\right] < \infty \quad \text{and} \quad \mathbb{E}_{\hat{t},\infty}\left[X_t(\cdot,x)^2\right] < \infty \quad \text{for all } x \in \mathbb{R}. \tag{2.8}$$

We will see later in the third step for which $t, \hat{t}$ and X_t these assumptions are fulfilled. One then shows that the FBSDE

$$\begin{aligned}
X_s &= X_t + \int_t^s \mu(r, X_r, Y_r, Z_r)\,\mathrm{d}r + \int_t^s \sigma(r, X_r, Y_r, Z_r)\,\mathrm{d}W_r, \\
Y_s &= \xi(X_T) - \int_s^T f(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_s^T Z_r\,\mathrm{d}W_r,
\end{aligned} \tag{2.9}$$

has a solution (X, Y, Z) on $[t,T]$ with initial condition X_t, where X, Y, Z are functions of (ω, s, x) and all equalities in (2.9) hold for all $x \in \mathbb{R}$. To that end, one chooses an arbitrary $x \in \mathbb{R}$, defines a suitable process space $(\mathcal{G}_{\hat{t}}, \|\cdot\|_w)$ and a mapping $F : \mathcal{G}_{\hat{t}} \to \mathcal{G}_{\hat{t}}$ such that every fixed point solves the above FBSDE. If $t \in [t', T)$ is close enough to T, i.e. t' is large enough, the function F is a contraction w.r.t. $\|\cdot\|_w$. Then according to Banach's fixed point theorem, the sequence $(X^k, Y^k, Z^k)_{k\in\mathbb{N}_0}$, recursively defined by

$$(X^k, Y^k, Z^k) := F(X^{k-1}, Y^{k-1}, Z^{k-1})$$

for $k \in \mathbb{N}$ and $(X^0, Y^0, Z^0) := (0, 0, 0)$, converges in $(\mathcal{G}_{\hat{t}}, \|\cdot\|_w)$ to a fixed point (X, Y, Z) that is unique. Consequently, (X, Y, Z) is the unique solution in $\mathcal{G}_{\hat{t}}$ to (2.9). Here the processes X, Y are unique up to modifications and Z is $\mathbb{P} \otimes \lambda$-almost everywhere unique. One can further show that (X^k, Y^k, Z^k) also converges in an L^2-sense and almost everywhere, and that X_s^k and Y_s^k converge almost surely to X_s and Y_s, respectively, for every fixed $s \in [t, T]$.

2. Show via induction over k that the processes X^k, Y^k, Z^k are progressively measurable and weakly differentiable w.r.t. x. If one chooses $t' < T$ large enough depending on $L, L_{\xi,x}, L_{\sigma,z}$, one can show that also the processes X, Y, Z are progressively measurable and weakly differentiable w.r.t. x, since (X^k, Y^k, Z^k) converges to (X, Y, Z).

3. Construct the decoupling field $u : \Omega \times [t'', T] \times \mathbb{R} \to \mathbb{R}$ as follows: For any $x \in \mathbb{R}$ and $t \in [t'', T]$ define

$$u(\cdot, t, x) := Y_t(\cdot, x),$$

where (X, Y, Z) is the unique solution in $\mathcal{G}_t$ to the FBSDE (2.9) with initial condition X_t given by $X_t(\omega, x) := x$. This solution exists because $\partial_x X_t = 1$ and $\mathbb{E}_{t,\infty}[X_t(\cdot, x)] < \infty$ for all $x \in \mathbb{R}$ and hence the requirements (2.8) in the first step are satisfied for $\hat{t} := t$. The time $t'' \in [0, T)$ is chosen such that $t'' \geq t'$ and large enough depending again on $L, L_{\xi,x}, L_{\sigma,z}$. One can then show that:

- For all $t \in [t'', T]$ it can be assumed that the mapping $u(\cdot, t, \cdot)$ is measurable and Lipschitz continuous in the last argument.

- It holds that $L_{u,x} < L_{\sigma,z}^{-1}$, $u(\cdot, \cdot, x)$ is progressively measurable for all $x \in \mathbb{R}$ and

$$\sup_{t \in [t'', T]} \|u(\cdot, t, 0)\|_\infty < \infty.$$

  As a consequence u is weakly regular.

4. Show that u is indeed a decoupling field by verifying the properties of Definition 2.9. In addition, one can show uniqueness of u.

In this summarized proof the choice of t' and t'' might appear arbitrary, but in the more detailed proof one can observe that the times $t', t'' < T$ with the necessary properties indeed exist. At last note that, for instance, the time $t = t''$ has the properties stated in Theorem 2.16. □

The construction above does not only provide the existence of a weakly regular decoupling field, but also yields:

- the uniqueness of the decoupling field up to modifications,

- the existence of a strongly regular decoupling field and

- the existence of a unique solution (X, Y, Z) to the FBSDE (2.6) on some time interval.

We summarize these properties later in Theorem 2.21. Here we mean by uniqueness that the processes X and Y are unique up to modifications and the process Z is $\mathbb{P} \otimes \lambda$-almost everywhere uniquely determined. We now turn to the consequences of Theorem 2.16, but omit the proofs. In particular, we present the statements of Corollary 2.5.3, Corollary 2.5.4 and Corollary 2.5.5 in [6].

Proposition 2.17. *Let (ξ, μ, σ, f) satisfy SLC.*

1. *If u_1 and u_2 are two weakly regular decoupling fields for (ξ, μ, σ, f) on some interval $[t, T]$, then u_1 and u_2 are modifications of each other.*
2. *If u is a weakly regular decoupling field for (ξ, μ, σ, f) on $[t, T]$, then u is strongly regular.*
3. *If u is a weakly regular decoupling field for (ξ, μ, σ, f) on $[t, T]$, then for any deterministic initial condition $X_t = x \in \mathbb{R}$ there is a unique solution (X, Y, Z) of the FBSDE (2.6) on $[t, T]$ such that*

$$\sup_{s \in [t,T]} \mathbb{E}\left[X_s^2\right] + \sup_{s \in [t,T]} \mathbb{E}\left[Y_s^2\right] + \mathbb{E}\left[\int_t^T Z_s^2 \, \mathrm{d}s\right] < \infty.$$

So far we only know that under SLC there is some time interval, on which we can construct a decoupling field and on which a solution exists. However, we would like to know how much we can enlarge this interval. This is because we actually want to solve the FBSDE on the whole interval $[0, T]$. Unfortunately, this is not always possible and therefore we define the *maximal interval* on which a weakly regular decoupling field exists as follows.

Definition 2.18. We say that $I_{max} \subseteq [0, T]$ is the *maximal interval* for (ξ, μ, σ, f) if it is the union of all intervals $[t, T] \subseteq [0, T]$ such that there exists a weakly regular decoupling field u on $[t, T]$.

Note that under SLC the interval I_{max} is either equal to $[0, T]$ or has the form $(t, T]$ for some $t \in (0, T)$. This result can be found in Theorem 2.5.11 of [6].

Lemma 2.19. *Let (ξ, μ, σ, f) satisfy SLC. Then either $I_{max} = [0, T]$ or $I_{max} = (t_{min}, T]$ holds true, where $0 \leq t_{min} < T$.*

Proof. We prove that $I_{max} = [t, T]$ for $t \in (0, T)$ is impossible. Therefore, we assume on the contrary that I_{max} has this form. Then there exists a weakly decoupling field u_1 on $[t, T]$. We can, however, use $u_1(t, \cdot)$ as terminal condition and apply Theorem 2.16. Thus we obtain a weakly regular decoupling field u_2 on the interval $[s, t]$ for some $s \in [0, t)$. Using Lemma 2.11 to concatenate u_1 and u_2 we get a decoupling field on the larger interval $[s, T]$, which contradicts our assumptions. $\square$

We now extend the notion of weak and strong regularity to half-open intervals.

Definition 2.20. Let $t \in [0, T)$ and $u : \Omega \times (t, T] \times \mathbb{R} \to \mathbb{R}$. We call u a decoupling field for (ξ, μ, σ, f) if for all $t' \in (t, T]$ the function $u\,|_{[t',T]}$ is a decoupling field for (ξ, μ, σ, f). Furthermore, we say that u is *weakly/strongly regular* if for all $t' \in (t, T]$ the function $u\,|_{[t',T]}$ is weakly/strongly regular.

By combining all the statements above we obtain the main result of this section (cf. Theorem 2.5.11 in [6]).

Theorem 2.21. *Let (ξ, μ, σ, f) satisfy SLC. Then there exists a unique weakly regular decoupling field u on I_{max}. This decoupling field is even strongly regular. It either holds $I_{max} = [0,T]$ or $I_{max} = (t_{min}, T]$, where $0 \leq t_{min} < T$. In addition, there is for every $t \in I_{max}$ and for any deterministic initial condition $X_t = x \in \mathbb{R}$ a unique solution $(X, Y, Z) \in \mathcal{S}^2_{t,T} \times \mathcal{S}^2_{t,T} \times \mathcal{H}^2_{t,T}$ to the FBSDE (2.6) on $[t,T]$.*

Sketch of the proof. We show how this result follows from statements presented in this section. By Theorem 2.16 there exists a weakly regular decoupling field u on I_{max}. Proposition 2.17 implies that u is strongly regular and by Lemma 2.19 it either holds $I_{max} = [0,T]$ or $I_{max} = (t_{min}, T]$, where $0 \leq t_{min} < T$. Moreover, there exists a solution (X, Y, Z) for the initial condition $X_t = x \in \mathbb{R}$ such that

$$\sup_{s\in[t,T]} \mathbb{E}\left[X_s^2\right] + \sup_{s\in[t,T]} \mathbb{E}\left[Y_s^2\right] + \mathbb{E}\left[\int_t^T Z_s^2 \,\mathrm{d}s\right] < \infty. \tag{2.10}$$

Here we mean by solution that for all $s \in [t,T]$ fixed the process (X,Y,Z) satisfies (2.6) almost surely. Furthermore, the processes X and Y are unique up to modifications and Z is $\mathbb{P} \otimes \lambda$-almost everywhere uniquely determined.

Note that it can be assumed that the processes X and Y are continuous in time since there are modifications of X and Y having this property. One can now show that the triple (X,Y,Z) satisfies the FBSDE (2.6) for all $s \in [t,T]$, $\mathbb{P}$-almost surely, using the continuity of X and Y. To be more precise, one can prove that there exists a $\mathbb{P}$-null set $N \in \mathcal{F}$ such that (2.6) is fulfilled for all $\omega \in N^c$ and $s \in [t,T]$.

It remains to show that $X, Y \in \mathcal{S}^2_{t,T}$, because then (X,Y,Z) is a solution to the FBSDE (2.6) in sense of Definition 2.6. We only prove that $X \in \mathcal{S}^2_{t,T}$. The statement for Y can be shown in the same way.

Note that, in particular, $X, Y, Z \in \mathcal{H}^2_{t,T}$. We observe that $M_\cdot := \int_t^\cdot \sigma(r, X_r, Y_r, Z_r)\,\mathrm{d}W_r$ is a martingale since

$$\begin{aligned}
&\mathbb{E}\left[\int_t^T |\sigma(r, X_r, Y_r, Z_r)|^2 \,\mathrm{d}r\right] \\
&\qquad \leq \mathbb{E}\left[\int_t^T \left(8L^2\left(X_r^2 + Y_r^2 + Z_r^2\right) + 2\|\sigma(\cdot,\cdot,0,0,0)\|_\infty^2\right) \mathrm{d}r\right] < \infty,
\end{aligned}$$

according to the Lipschitz continuity of σ. As aforementioned, the process X satisfies the forward equation in the FBSDE (2.6) for all $s \in [t,T]$, $\mathbb{P}$-almost surely, and thus we have

$$\begin{aligned}
\mathbb{E}\left[\sup_{s\in[t,T]} X_s^2\right] &\leq 4x^2 + 4\,\mathbb{E}\left[\sup_{s\in[t,T]} M_s^2\right] + 4\,\mathbb{E}\left[\sup_{s\in[t,T]}\left(\int_t^s \mu(r, X_r, Y_r, Z_r)\,\mathrm{d}r\right)^2\right] \\
&\leq 4x^2 + 16\,\mathbb{E}\left[M_T^2\right] + 4\,\mathbb{E}\left[\sup_{s\in[t,T]}\left((s-t)\int_t^s |\mu(r, X_r, Y_r, Z_r)|^2\,\mathrm{d}r\right)\right] \\
&\leq 4x^2 + 16\,\mathbb{E}\left[\int_t^T |\sigma(r, X_r, Y_r, Z_r)|^2\,\mathrm{d}r\right] \\
&\qquad + 4T\,\mathbb{E}\left[\int_t^T \left(8L^2\left(X_r^2 + Y_r^2 + Z_r^2\right) + 2\|\mu(\cdot,\cdot,0,0,0)\|_\infty^2\right) \mathrm{d}r\right] < \infty.
\end{aligned}$$

We have used Doob's L^2-inequality, Jensen's inequality, the Itô isometry and the Lipschitz continuity of μ. $\square$

If the maximal interval I_{max} is equal to $(t_{min}, T]$, we have the following significant property that allows us to develop the method of decoupling fields (cf. Lemma 2.5.12 in [6]).

Proposition 2.22. *Let (ξ, μ, σ, f) satisfy SLC, $I_{max} = (t_{min}, T]$ and u be the unique decoupling field on I_{max}. Then we have that*

$$\lim_{t \downarrow t_{min}} L_{u(t,\cdot),x} = L_{\sigma,z}^{-1}. \tag{2.11}$$

Under the assumption that the parameters satisfy SLC, this proposition allows us to formulate an algorithm for checking if an FBSDE has a global solution, i.e. a solution on the whole interval $[0, T]$. It is based on the idea that if (2.11) does not hold true, $I_{max} = [0, T]$ is the only possible choice according to Theorem 2.21 and Proposition 2.22. We refer to this technique as the *method of decoupling fields*. In more detail, the algorithm is given by:

1. Assume that $I_{max} = (t_{min}, T]$ holds for $0 \leq t_{min} < T$. Then there exists a strongly regular decoupling field u on I_{max}. We fix an arbitrary $t \in I_{max}$ and $x \in \mathbb{R}$. According to Theorem 2.21 there exists a solution (X, Y, Z) to the FBSDE on $[t, T]$ with initial condition $X_t = x$ that is weakly differentiable w.r.t. the initial value x.

2. Differentiate the FBSDE w.r.t. x in the weak sense, which is possible due to the strong regularity of u. This yields an FBSDE for $\partial_x X, \partial_x Y, \partial_x Z$.

3. Analyse the dynamics of $\partial_x Y_s(\partial_x X_s)^{-1}$ with the help of Itô's formula. One expects that $\partial_x Y_s(\partial_x X_s)^{-1} = u_x(s, X_s)$ holds as a consequence of differentiating the decoupling condition $Y_s = u(s, X_s)$ using the chain rule in Proposition 2.38.

4. Use the dynamics of $u_x(s, X_s)$ to show that $L_{u(t,\cdot),x}$ can be bounded away from $L_{\sigma,z}^{-1}$ independent of t, x, s, ω. This contradicts Proposition 2.22 and hence only $I_{max} = [0, T]$ can be true.

In chapter 3 we follow these steps to obtain the global existence of a solution to a certain FBSDE derived from the maximum principle, which we introduce in the following section. All the above considerations are especially developed for the non-Markovian case. We emphasize that in the Markovian case, i.e. if μ, σ, f, ξ are deterministic, one can lessen the requirement of SLC. It suffices to consider certain local Lipschitz conditions. Moreover, one can show that under certain assumptions the decoupling field is also deterministic and even continuous (see e.g. Theorem 2.5.18 in [6]).

We conclude this chapter with an example.

Example 2.23. We extend Example 2.8. Recall that we considered the FBSDE on $[t, T]$

$$\begin{aligned} X_s &= x + \int_t^s Y_r \, \mathrm{d}r, \\ Y_s &= X_T - \int_s^T Z_r \, \mathrm{d}W_r, \ \ s \in [t, T] \end{aligned} \tag{2.12}$$

for $T = 1,\ t \in [0, T],\ x \in \mathbb{R}$. It was possible to find a solution to the above FBSDE if either $x = 0$ or $t \in (0, T]$. In the case of $x \neq 0$ and $t = 0$ we showed that there is no solution.

At first, note that $L_{\xi,x} = 1, L_{\sigma,z}^{-1} = \infty$ (since $L_{\sigma,z} = 0$) and thus the parameters of the FBSDE (2.12) satisfy SLC. We want to show how a decoupling field looks like in the case of $t \in (0, T]$. The above problem has the solution

$$X_s = x\frac{s}{t}, \quad Y_s = \frac{x}{t}, \quad Z_s = 0,\ s \in [t, T],$$

according to Example 2.8. A decoupling field is therefore given by

$$u(s, \hat{x}) := \frac{\hat{x}}{s},\ s \in [t, T],\ \hat{x} \in \mathbb{R}.$$

Moreover, we observe that $L_{u(s,\cdot),x} = \frac{1}{s}$ for all $s \in [t, T]$, $L_{u,x} = \frac{1}{t} < L_{\sigma,z}^{-1} = \infty$ and $\sup_{s\in[t,T]} |u(s, 0)| < \infty$. Consequently, the decoupling field u is weakly regular and we have

$$\lim_{s\downarrow 0} L_{u(s,\cdot),x} = \lim_{s\downarrow 0} \frac{1}{s} = \infty = L_{\sigma,z}^{-1}.$$

Hence in this example the statement of Proposition 2.22 is fulfilled. However, this does not prove that $I_{max} \neq [0, T]$. Note that $\lim_{t\downarrow 0} u(t, \hat{x}) = \infty$ for $\hat{x} \neq 0$, which suggests that the problem might not have a solution on the whole interval $[0, T]$. In addition, the decoupling field has to be continuous on $[0, T]$ for a global solution as stated in Example 2.3.2 of [6]. This is obviously not the case, since u cannot be continuously extended to $[0, T]$.

2.4 Pontryagin's maximum principle

The maximum principle, introduced by Pontryagin amongst others, gives a powerful technique to solve optimal control problems. In contrast to the HJB approach, where one has to solve a PDE analytically or numerically, the maximum principle transforms the optimization problem into the task of solving a possibly coupled FBSDE. As we have seen in the previous sections, a coupled FBSDE might not be easy to solve or even be unsolvable. However, this method can be more suitable for a certain class of problems, for instance, if the functions that describe the running and terminal costs depend on ω.

This section is based on chapter 6.4.2 of [11]. In contrast to the book by Pham, who maximizes the target function, we consider the problem of minimizing it because of our application of this theory in chapter 3, where also a minimization problem is studied. Consequently, we have to adjust the requirements of the maximum principle. We assume convexity of the function g and of the Hamiltonian $\mathcal{H}$ instead of concavity. Moreover, we study, unlike Pham, a more general case in allowing ω-dependent functions μ, σ, f, g. As a result of our different framework, the Hamiltonian and the adjoint equation differ slightly from the ones considered by Pham. The formulation of the maximum principle and its proof, however, remain very similar.

Before we turn to the statement of Pontryagin's maximum principle, we introduce the following. Let $A \subseteq \mathbb{R}$ be measurable and non-empty. Let $\mu, \sigma : \Omega \times [0, T] \times \mathbb{R} \times A \to \mathbb{R}$

be measurable functions that denote the drift and diffusion coefficient, respectively. Moreover, we define the measurable functions $f : \Omega \times [0,T] \times \mathbb{R} \times A \to \mathbb{R}$ and $g : \Omega \times \mathbb{R} \to \mathbb{R}$. We assume:

(A1) The mappings μ, σ, f are progressively measurable in the sense that for all $t \in [0,T]$ the restriction of μ, σ to $\Omega \times [0,t] \times \mathbb{R} \times \mathbb{R}$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t]) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$-measurable and the restriction of f to $\Omega \times [0,t] \times \mathbb{R} \times A$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t]) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(A)$-measurable.

(A2) μ and σ satisfy a Lipschitz condition, i.e. there exists a constant $L \geq 0$ such that for all $t \in [0,T]$, $x, y \in \mathbb{R}$, $a \in A$ and $\omega \in \Omega$

$$|\mu(\omega,t,x,a) - \mu(\omega,t,y,a)| + |\sigma(\omega,t,x,a) - \sigma(\omega,t,y,a)| \leq L|x-y|.$$

Moreover, it holds for all $a \in A$ that

$$\mathbb{E}\left[\int_0^T |\mu(t,0,a)|^2 + |\sigma(t,0,a)|^2 \,\mathrm{d}t\right] < \infty.$$

(A3) For all $\omega \in \Omega$, $a \in A$ fixed we have that $f(\omega,\cdot,\cdot,a)$ is continuous and $g(\omega,\cdot)$ is convex and continuously differentiable. Additionally, f and g satisfy a quadratic growth condition in x, i.e. there exists a constant $C > 0$ such that for $\mathbb{P} \otimes \lambda$-almost every $(\omega,t) \in \Omega \times [0,T]$ it holds for all $x \in \mathbb{R}, a \in A$

$$\begin{aligned} |f(\omega,t,x,a)| &\leq C\left(1 + x^2 + |\mu(\omega,t,0,a)|^2 + |\sigma(\omega,t,0,a)|^2\right), \text{ and} \\ |g(\omega,x)| &\leq C\left(1 + x^2\right). \end{aligned}$$

(A4) The *Hamiltonian* $\mathcal{H} : \Omega \times [0,T] \times \mathbb{R} \times A \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, defined by

$$\mathcal{H}(\omega,t,x,a,y,z) := \mu(\omega,t,x,a)y + \sigma(\omega,t,x,a)z + f(\omega,t,x,a),$$

is differentiable in (x,a) for all $(\omega,t,y,z) \in \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}$.

Let $\mathcal{A}$ denote the set of all *admissible controls*, i.e. $\mathcal{A}$ is the set of all progressively measurable control processes $\alpha : \Omega \times [0,T] \to A$ such that

$$\mathbb{E}\left[\int_0^T |\mu(t,0,\alpha_t)|^2 + |\sigma(t,0,\alpha_t)|^2 \,\mathrm{d}t\right] < \infty.$$

Note that $\mathcal{A}$ is non-empty since is contains at least all constant controls due to (A2).

For the initial value $x \in \mathbb{R}$ and the control $\alpha \in \mathcal{A}$ we consider the controlled diffusion process $X^{x,\alpha}$, driven by the SDE

$$X_t^{x,\alpha} = x + \int_0^t \mu(s, X_s^{x,\alpha}, \alpha_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s^{x,\alpha}, \alpha_s)\,\mathrm{d}W_s, \;\; t \in [0,T]. \tag{2.13}$$

From assumptions (A1), (A2) and the definition of $\mathcal{A}$ we gather that there indeed exists a unique solution to (2.13) (see e.g. Theorem 1.3.15 in [11]). We study the target function J, given by

$$J(x,\alpha) := \mathbb{E}\left[\int_0^T f(t, X_t^{x,\alpha}, \alpha_t)\,\mathrm{d}t + g(X_T^{x,\alpha})\right],$$

that can be interpreted as the total costs of the control problem, because we think of the functions f and g as running and terminal costs, respectively. As we want to lower the costs, our goal is to find a control $\hat{\alpha} \in \mathcal{A}$ such that

$$J(x, \hat{\alpha}) = \inf_{\alpha \in \mathcal{A}} J(x, \alpha). \tag{2.14}$$

We call such an $\hat{\alpha}$ an *optimal control*.

Remark 2.24. The target function J is well-defined since by (A2) and (A3) we observe that

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T |f(t, X_t^{x,\alpha}, \alpha_t)| \,\mathrm{d}t + |g(X_T^{x,\alpha})|\right] \\
&\quad \leq \mathbb{E}\left[\int_0^T C\left(1 + (X_t^{x,\alpha})^2 + |\mu(t, 0, \alpha_t)|^2 + |\sigma(t, 0, \alpha_t)|^2\right) \,\mathrm{d}t + C\left(1 + (X_T^{x,\alpha})^2\right)\right] < \infty.
\end{aligned}$$

The last term is finite, because $\alpha \in \mathcal{A}$ and $X^{x,\alpha} \in \mathcal{S}^2_{0,T}$ (cf. Theorem 1.3.15 in [11]).

The essential point in Pontryagin's maximum principle is to consider the so-called *adjoint equation*, given by

$$\begin{aligned}
X_t^{x,\alpha} &= x + \int_0^t \mu(s, X_s^{x,\alpha}, \alpha_s) \,\mathrm{d}s + \int_0^t \sigma(s, X_s^{x,\alpha}, \alpha_s) \,\mathrm{d}W_s, \\
Y_t^{x,\alpha} &= g'(X_T^{x,\alpha}) + \int_t^T D_x\mathcal{H}(s, X_s^{x,\alpha}, \alpha_s, Y_s^{x,\alpha}, Z_s^{x,\alpha}) \,\mathrm{d}s - \int_t^T Z_s^{x,\alpha} \,\mathrm{d}W_s, \;\; t \in [0, T],
\end{aligned} \tag{2.15}$$

to determine an optimal control. This equation is an FBSDE that might be coupled, depending on the particular choice of α. Hence there does not have to exist a solution as explained in the recent sections. Assuming, however, that there is a solution, the following theorem gives sufficient conditions for the optimality of a control.

From now on we omit the initial value x and the control α in the superscript of the processes X, Y, Z to simplify notation. Nevertheless, one should always keep this dependence in mind and therefore we will always mention the initial value x and the control α we are considering.

Theorem 2.25 (Pontryagin's maximum principle)**.** *Let the assumptions mentioned above hold true, $\hat{\alpha} \in \mathcal{A}$, $x_0 \in \mathbb{R}$ and $(\hat{X}, \hat{Y}, \hat{Z})$ be a solution to the FBSDE (2.15) with initial value x_0 and control $\hat{\alpha}$. For $\mathbb{P} \otimes \lambda$-almost every $(\omega, t) \in \Omega \times [0, T]$ we assume that*

$$\mathcal{H}(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t) = \min_{a \in A} \mathcal{H}(t, \hat{X}_t, a, \hat{Y}_t, \hat{Z}_t), \tag{2.16}$$

and that the mapping

$$(x, a) \mapsto \mathcal{H}(t, x, a, \hat{Y}_t, \hat{Z}_t) \text{ is convex.} \tag{2.17}$$

Then $\hat{\alpha}$ is an optimal control, i.e. $J(x_0, \hat{\alpha}) = \inf_{\alpha \in \mathcal{A}} J(x_0, \alpha)$.

Proof. We adapt the proof of Theorem 6.4.6 in [11] to our setting. Let $\alpha \in \mathcal{A}$ and X be a solution to the SDE (2.13) with initial value x_0 and control α. In particular, this means that $X \in \mathcal{S}^2_{0,T}$. Moreover, we have $(\hat{X}, \hat{Y}, \hat{Z}) \in \mathcal{S}^2_{0,T} \times \mathcal{S}^2_{0,T} \times \mathcal{H}^2_{0,T}$ since it is a solution to the FBSDE (2.15). We aim at showing that $J(x_0, \hat{\alpha}) \leq J(x_0, \alpha)$ and therefore we consider

$$J(x_0, \alpha) - J(x_0, \hat{\alpha}) = \mathbb{E}\left[\int_0^T \Big(f(t, X_t, \alpha_t) - f(t, \hat{X}_t, \hat{\alpha}_t)\Big)\, \mathrm{d}t + g(X_T) - g(\hat{X}_T)\right].$$

We calculate the two terms in the expectation and show that the right hand side is larger or equal to zero. By the convexity of g (cf. (A3)) we observe that

$$g(X_T) - g(\hat{X}_T) \geq g'(\hat{X}_T)(X_T - \hat{X}_T) = \hat{Y}_T(X_T - \hat{X}_T), \text{ a.s.} \tag{2.18}$$

In order to apply the product formula to $\hat{Y}(X - \hat{X})$, we write $\hat{Y}$ as an Itô process

$$\hat{Y}_t = \hat{Y}_0 - \int_0^t D_x\mathcal{H}(s, \hat{X}_s, \hat{\alpha}_s, \hat{Y}_s, \hat{Z}_s)\, \mathrm{d}s + \int_0^t \hat{Z}_s\, \mathrm{d}W_s,\ t \in [0, T],$$

and consequently we obtain that

$$\begin{aligned}
&\hat{Y}_t(X_t - \hat{X}_t)\\
&= \int_0^t (X_s - \hat{X}_s)\, \mathrm{d}\hat{Y}_s + \int_0^t \hat{Y}_s\, \mathrm{d}(X_s - \hat{X}_s) + \int_0^t \Big(\sigma(s, X_s, \alpha_s) - \sigma(s, \hat{X}_s, \hat{\alpha}_s)\Big) \hat{Z}_s\, \mathrm{d}s\\
&= -\int_0^t (X_s - \hat{X}_s) D_x\mathcal{H}(s, \hat{X}_s, \hat{\alpha}_s, \hat{Y}_s, \hat{Z}_s)\, \mathrm{d}s + \int_0^t \hat{Y}_s \Big(\mu(s, X_s, \alpha_s) - \mu(s, \hat{X}_s, \hat{\alpha}_s)\Big)\, \mathrm{d}s\\
&\qquad + \int_0^t \Big(\sigma(s, X_s, \alpha_s) - \sigma(s, \hat{X}_s, \hat{\alpha}_s)\Big) \hat{Z}_s\, \mathrm{d}s\\
&\qquad\quad + \int_0^t \Big((X_s - \hat{X}_s)\hat{Z}_s + \hat{Y}_s \Big(\sigma(s, X_s, \alpha_s) - \sigma(s, \hat{X}_s, \hat{\alpha}_s)\Big)\Big)\, \mathrm{d}W_s,
\end{aligned} \tag{2.19}$$

for all $t \in [0, T]$, $\mathbb{P}$-almost surely. Furthermore, the definition of $\mathcal{H}$ allows us to rewrite the integral below as follows:

$$\begin{aligned}
&\int_0^t \Big(f(s, X_s, \alpha_s) - f(s, \hat{X}_s, \hat{\alpha}_s)\Big)\, \mathrm{d}s\\
&\qquad = \int_0^t \Big(\mathcal{H}(s, X_s, \alpha_s, \hat{Y}_s, \hat{Z}_s) - \mathcal{H}(s, \hat{X}_s, \hat{\alpha}_s, \hat{Y}_s, \hat{Z}_s)\Big)\, \mathrm{d}s\\
&\qquad\quad - \int_0^t \Big(\mu(s, X_s, \alpha_s) - \mu(s, \hat{X}_s, \hat{\alpha}_s)\Big) \hat{Y}_s\, \mathrm{d}s\\
&\qquad\qquad - \int_0^t \Big(\sigma(s, X_s, \alpha_s) - \sigma(s, \hat{X}_s, \hat{\alpha}_s)\Big) \hat{Z}_s\, \mathrm{d}s,\ t \in [0, T], \text{ a.s.}
\end{aligned} \tag{2.20}$$

By (2.19) and (2.20) we observe that $\mathbb{P}$-almost surely for all $t \in [0, T]$

$$\begin{aligned}
&\int_0^t \Big(f(s, X_s, \alpha_s) - f(s, \hat{X}_s, \hat{\alpha}_s)\Big)\, \mathrm{d}s + \hat{Y}_t(X_t - \hat{X}_t)\\
&\qquad \geq \int_0^t \Big(\mathcal{H}(s, X_s, \alpha_s, \hat{Y}_s, \hat{Z}_s) - \mathcal{H}(s, \hat{X}_s, \hat{\alpha}_s, \hat{Y}_s, \hat{Z}_s)\\
&\qquad\quad - (X_s - \hat{X}_s) D_x\mathcal{H}(s, \hat{X}_s, \hat{\alpha}_s, \hat{Y}_s, \hat{Z}_s)\Big)\, \mathrm{d}s\\
&\qquad\qquad + \int_0^t \Big((X_s - \hat{X}_s)\hat{Z}_s + \hat{Y}_s \Big(\sigma(s, X_s, \alpha_s) - \sigma(s, \hat{X}_s, \hat{\alpha}_s)\Big)\Big)\, \mathrm{d}W_s.
\end{aligned} \tag{2.21}$$

Now we fix $(\omega, s) \in \Omega \times [0,T]$ such that (2.16) and (2.17) hold true. We observe that by the convexity of $(x,a) \mapsto \mathcal{H}(t,x,a,\hat{Y}_t,\hat{Z}_t)$ and the differentiability of $\mathcal{H}$ in (x,a) (cf. (A4))

$$
\begin{aligned}
&\mathcal{H}(s,X_s,\alpha_s,\hat{Y}_s,\hat{Z}_s) - \mathcal{H}(s,\hat{X}_s,\hat{\alpha}_s,\hat{Y}_s,\hat{Z}_s) \\
&\geq (X_s-\hat{X}_s)D_x\mathcal{H}(s,\hat{X}_s,\hat{\alpha}_s,\hat{Y}_s,\hat{Z}_s) + (\alpha_s-\hat{\alpha}_s)D_a\mathcal{H}(s,\hat{X}_s,\hat{\alpha}_s,\hat{Y}_s,\hat{Z}_s) \\
&\geq (X_s-\hat{X}_s)D_x\mathcal{H}(s,\hat{X}_s,\hat{\alpha}_s,\hat{Y}_s,\hat{Z}_s),
\end{aligned}
$$

where $(\alpha_s-\hat{\alpha}_s)D_a\mathcal{H}(s,\hat{X}_s,\hat{\alpha}_s,\hat{Y}_s,\hat{Z}_s) \geq 0$ by (2.16). This identity is true for $\mathbb{P}\otimes\lambda$-almost every $(\omega,s) \in \Omega \times [0,T]$. Thus, by plugging this estimate into (2.21) we obtain

$$
\begin{aligned}
&\int_0^t \Big(f(s,X_s,\alpha_s) - f(s,\hat{X}_s,\hat{\alpha}_s)\Big)\,\mathrm{d}s + \hat{Y}_t(X_t-\hat{X}_t) \\
&\geq \int_0^t \Big((X_s-\hat{X}_s)\hat{Z}_s + \hat{Y}_s\Big(\sigma(s,X_s,\alpha_s)-\sigma(s,\hat{X}_s,\hat{\alpha}_s)\Big)\Big)\,\mathrm{d}W_s,\ t\in[0,T],\ \text{a.s.}
\end{aligned}
\tag{2.22}
$$

Taking the expectation on both sides for $t = T$ provides the result. However, in more detail we have to show that the expectation of the right hand side is indeed equal to zero. Let $(\tau_n)_{n\in\mathbb{N}}$ be a localizing sequence, defined by

$$
\tau_n := \inf\{t \geq 0 : |X_t - \hat{X}_t| \geq n \text{ or } |\hat{Y}_t| \geq n\} \wedge T,\ n \in \mathbb{N}.
$$

We observe that $\tau_n \to T$ a.s. as $n \to \infty$, because $(X-\hat{X}),\hat{Y} \in \mathcal{S}^2_{0,T}$. Moreover, we have for all $n \in \mathbb{N}$

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^{\tau_n}\Big((X_s-\hat{X}_s)\hat{Z}_s + \hat{Y}_s\Big(\sigma(s,X_s,\alpha_s)-\sigma(s,\hat{X}_s,\hat{\alpha}_s)\Big)\Big)^2\,\mathrm{d}s\right] \\
&\quad\leq \mathbb{E}\left[\int_0^{\tau_n}\Big(2(X_s-\hat{X}_s)^2\hat{Z}_s^2 + 2\hat{Y}_s^2\Big(\sigma(s,X_s,\alpha_s)-\sigma(s,\hat{X}_s,\hat{\alpha}_s)\Big)^2\Big)\,\mathrm{d}s\right] \\
&\quad\leq \mathbb{E}\left[4n^2\int_0^T\Big(\hat{Z}_s^2 + \sigma(s,X_s,\alpha_s)^2 + \sigma(s,\hat{X}_s,\hat{\alpha}_s)^2\Big)\,\mathrm{d}s\right] \\
&\quad\leq \mathbb{E}\left[4n^2\int_0^T\Big(\hat{Z}_s^2 + 2L^2\Big(X_s^2+\hat{X}_s^2\Big) + 2\sigma(s,0,\alpha_s)^2 + 2\sigma(s,0,\hat{\alpha}_s)^2\Big)\,\mathrm{d}s\right] < \infty,
\end{aligned}
$$

where we have used the Lipschitz continuity of σ. The last term is finite because of $X,\hat{X} \in \mathcal{S}^2_{0,T}$, $\hat{Z} \in \mathcal{H}^2_{0,T}$ and $\alpha,\hat{\alpha}$ being admissible controls. Consequently, the process M, defined by

$$
M_t := \int_0^{\tau_n\wedge t}\Big((X_s-\hat{X}_s)\hat{Z}_s + Y_s\Big(\sigma(s,X_s,\alpha_s)-\sigma(s,\hat{X}_s,\hat{\alpha}_s)\Big)\Big)\,\mathrm{d}W_s,\ t\in[0,T],
$$

is a martingale on $[0,T]$ for all fixed $n \in \mathbb{N}$ and therefore its expectation vanishes. Thus, we have for all $n \in \mathbb{N}$

$$
\mathbb{E}\left[\int_0^{\tau_n}\Big(f(s,X_s,\alpha_s)-f(s,\hat{X}_s,\hat{\alpha}_s)\Big)\,\mathrm{d}s + \hat{Y}_{\tau_n}(X_{\tau_n}-\hat{X}_{\tau_n})\right] \geq 0. \tag{2.23}
$$

Note that the expectation exists due to $X, \hat{X}, \hat{Y} \in \mathcal{S}^2_{0,T}$ and Remark 2.24. As a matter of fact, it even exists for time T. Therefore, dominated convergence and (2.18) imply

$$\begin{aligned} J(x_0,\alpha) - J(x_0,\hat{\alpha}) &= \mathbb{E}\left[\int_0^T \Big(f(s,X_s,\alpha_s) - f(s,\hat{X}_s,\hat{\alpha}_s)\Big)\,\mathrm{d}s + g(X_T) - g(\hat{X}_T)\right] \\ &\geq \mathbb{E}\left[\int_0^T \Big(f(s,X_s,\alpha_s) - f(s,\hat{X}_s,\hat{\alpha}_s)\Big)\,\mathrm{d}s + \hat{Y}_T(X_T - \hat{X}_T)\right] \\ &= \lim_{n\to\infty} \mathbb{E}\left[\int_0^{\tau_n} \Big(f(s,X_s,\alpha_s) - f(s,\hat{X}_s,\hat{\alpha}_s)\Big)\,\mathrm{d}s + \hat{Y}_{\tau_n}(X_{\tau_n} - \hat{X}_{\tau_n})\right] \\ &\geq 0 \end{aligned}$$

and hence $J(x_0,\hat{\alpha}) \leq J(x_0,\alpha)$. We emphasize that it is straightforward to find a dominating function for (2.23) using again the fact that $X, \hat{X}, \hat{Y} \in \mathcal{S}^2_{0,T}$ and the estimate in Remark 2.24. This concludes the proof. □

Remark 2.26. We point out that due to the dependence on the control α it is in general not clear if there exists a solution (X,Y,Z) to (2.15). However, one can try to solve the minimization problem (2.14) by following the algorithm:

1. Determine $a_* = a_*(\omega,t,x,y,z)$ such that

$$\mathcal{H}(\omega,t,x,a_*,y,z) = \min_{a\in A} \mathcal{H}(\omega,t,x,a,y,z).$$

2. Check if the FBSDE (2.15), controlled by $\alpha_t(\omega) := a_*(\omega,t,X_t,Y_t,Z_t),\ t\in[0,T]$, has a solution (X,Y,Z).

3. Apply Theorem 2.25, if possible, to obtain that α is an optimal control.

In section 3 we follow these steps to determine an optimal control. We will see that the second step is the most challenging one.

We conclude this section with a generalization of the maximum principle. We consider the FBSDE (2.15) starting at an arbitrary time $t\in[0,T]$ instead of starting in 0, i.e. we consider the adjoint FBSDE

$$\begin{aligned} X_s^{t,x,\alpha} &= x + \int_t^s \mu(r,X_r^{t,x,\alpha},\alpha_r)\,\mathrm{d}r + \int_t^s \sigma(r,X_r^{t,x,\alpha},\alpha_r)\,\mathrm{d}W_r, \\ Y_s^{t,x,\alpha} &= g'(X_T^{t,x,\alpha}) + \int_s^T D_x\mathcal{H}(r,X_r^{t,x,\alpha},\alpha_r,Y_r^{t,x,\alpha},Z_r^{t,x,\alpha})\,\mathrm{d}r - \int_s^T Z_r^{t,x,\alpha}\,\mathrm{d}W_r, \end{aligned} \tag{2.24}$$

for $s\in[t,T], x\in\mathbb{R}, \alpha\in\mathcal{A}(t)$. The control space $\mathcal{A}(t)$ is an adaptation of $\mathcal{A}$ to this particular setting and is defined as the set of all progressively measurable control processes $\alpha : [t,T]\times\Omega\to A$ such that

$$\mathbb{E}\left[\int_t^T |\mu(t,0,\alpha_t)|^2 + |\sigma(t,0,\alpha_t)|^2\,\mathrm{d}t\right] < \infty.$$

In this setting the target function is given by

$$J(\omega,t,x,\alpha) := \mathbb{E}\left[\int_t^T f(s,X_s^{t,x,\alpha},\alpha)\,\mathrm{d}s + g(X_T^{t,x,\alpha})\middle|\mathcal{F}_t\right](\omega),$$

for $(\omega, t, x) \in \Omega \times [0, T] \times \mathbb{R}$ and $\alpha \in \mathcal{A}(t)$. Moreover, we introduce the value function $v : \Omega \times [0, T] \times \mathbb{R} \to [-\infty, \infty]$ that is defined by

$$v(t, x) := \operatorname*{ess\,inf}_{\alpha \in \mathcal{A}(t)} J(t, x, \alpha).$$

To be more precise, $v(t, x)$ is defined to be an $\mathcal{F}_t$-measurable random variable satisfying:

(1) $v(t, x) \leq J(t, x, \alpha)$ a.s. for all $\alpha \in \mathcal{A}(t)$,

(2) $v(t, x) \geq V$ a.s., for each $\mathcal{F}_t$-measurable $V : \Omega \to [-\infty, \infty]$ having the property $V \leq J(t, x, \alpha)$ a.s. for all $\alpha \in \mathcal{A}(t)$.

Remark 2.27. The property (2) immediately implies the uniqueness of an $\mathcal{F}_t$-measurable random variable satisfying (1) and (2). In addition, one can show that such a random variable indeed exists. For fixed (t, x) there is even a countable subset $\mathcal{J} \subseteq \mathcal{A}(t)$ such that

$$v(\omega, t, x) = \inf_{\alpha \in \mathcal{J}} J(\omega, t, x, \alpha)$$

for almost all $\omega \in \Omega$, i.e. one can replace the essential infimum by the pointwise infimum. Note that the definition of the essential infimum over a family of random variables and the aforementioned properties can be found, for instance, in chapter 1.2 of [2] or in chapter A.2 of [11].

The choice of α ensures that for all $(t, x) \in [0, T] \times \mathbb{R}$ there exists a solution to the forward equation in (2.24), and that the target function is well-defined as in the first part of this section (see also Remark 2.24). We are interested in the case where there is for $(t, x) \in [0, T] \times \mathbb{R}$ a control $\hat{\alpha} \in \mathcal{A}(t)$ such that $v(t, x) = J(t, x, \hat{\alpha})$ a.s., i.e. there is an optimal control. The maximum principle in the generalized form gives sufficient conditions under which a control is optimal.

Theorem 2.28. *Let the assumptions of this section hold true,* $(t, x) \in [0, T] \times \mathbb{R}, \hat{\alpha} \in \mathcal{A}(t)$ *and* $(\hat{X}^{t,x,\hat{\alpha}}, \hat{Y}^{t,x,\hat{\alpha}}, \hat{Z}^{t,x,\hat{\alpha}})$ *be a solution to the FBSDE* (2.24). *For* $\mathbb{P} \otimes \lambda$*-almost every* $(\omega, s) \in \Omega \times [t, T]$ *we assume that*

$$\mathcal{H}(s, \hat{X}_s^{t,x,\hat{\alpha}}, \hat{\alpha}_s, \hat{Y}_s^{t,x,\hat{\alpha}}, \hat{Z}_s^{t,x,\hat{\alpha}}) = \min_{a \in A} \mathcal{H}(s, \hat{X}_s^{t,x,\hat{\alpha}}, a, \hat{Y}_s^{t,x,\hat{\alpha}}, \hat{Z}_s^{t,x,\hat{\alpha}}), \tag{2.25}$$

and that the mapping

$$(x', a) \mapsto \mathcal{H}(s, x', a, \hat{Y}_s^{t,x,\hat{\alpha}}, \hat{Z}_s^{t,x,\hat{\alpha}}) \text{ is convex.} \tag{2.26}$$

Then $\hat{\alpha}$ *is an optimal control, i.e.* $J(t, x, \hat{\alpha}) = v(t, x)$ *a.s.*

Sketch of the proof. We basically just generalize the proof of Theorem 2.25. Let $\alpha \in \mathcal{A}(t)$ and $X^{t,x,\alpha}$ be a solution to the forward equation in (2.24) controlled by α and starting at time t in x. Then by following the same train of thought, but considering the time interval $[t, T]$ instead of $[0, T]$ and replacing the expectation with the conditional one, we obtain

$$J(t, x, \hat{\alpha}) \leq J(t, x, \alpha), \text{ a.s.}$$

Hence we have according to the definition of the value function v that on the one hand $v(t, x) \leq J(t, x, \hat{\alpha})$ a.s. by (1), and on the other hand $v(t, x) \geq J(t, x, \hat{\alpha})$ a.s. by (2). Finally, $v(t, x) = J(t, x, \hat{\alpha})$ a.s. and $\hat{\alpha}$ is optimal. $\square$

2.5 Weak derivatives

In this thesis we often work with weak differentiability instead of the classical one, because we are focused on the method of decoupling fields, where in general only weak differentiability is provided. Furthermore, a benefit in considering weak derivatives is that one does not have to deal with differential quotients and their convergence. The weak derivative merely has to satisfy the integral equation (2.27).

This section introduces weak derivatives and states some important properties, which will be of importance later on in chapter 3 and 4. We consider only the one-dimensional case, but emphasize that all the statements have multidimensional analogues. Furthermore, we point out that the following is based on [6] and [14], i.e. one can find all the definitions and statements in these works.

Now we define for an open set $A \subseteq \mathbb{R}$ the space of all locally integrable functions defined on A by

$$\mathcal{L}^1_{loc}(A) := \left\{ f : A \to \mathbb{R} \text{ measurable} \,\middle|\, \forall K \subseteq A \text{ compact} : \int_K |f(x)| \, \mathrm{d}x < \infty \right\}.$$

In the usual way, we denote by $L^1_{loc}(A)$ the quotient space of $\mathcal{L}^1_{loc}(A)$ w.r.t. the equivalence relation, where two functions are equivalent if they coincide almost everywhere in the Lebesgue sense. We call such equivalent functions also *versions* of each other. Weak differentiability is defined as follows.

Definition 2.29. Let $A \subseteq \mathbb{R}$ be open and $f \in L^1_{loc}(A)$. We say that f is *weakly differentiable* if there exists a mapping $g \in L^1_{loc}(A)$ such that

$$\int_A \varphi(x) g(x) \, \mathrm{d}x = -\int_A \varphi'(x) f(x) \, \mathrm{d}x \tag{2.27}$$

for all $\varphi \in C_c^\infty(A)$. We denote the function g by $\partial_x f$ and call it the weak derivative of f.

Furthermore, let $(M, \mathcal{G}, \mu)$ be a finite measure space. We call a measurable function $X : M \times \mathbb{R} \to \mathbb{R}$ *weakly differentiable w.r.t.* x if the function $X(\omega, \cdot)$ is weakly differentiable for μ-almost every $\omega \in M$. We denote by $\partial_x X(\omega, x)$ the weak derivative w.r.t. x.

Remark 2.30. Concerning the definition we mention:

- The second part of the definition is crucial since we mainly calculate weak derivatives of stochastic processes or random variables. In these cases we have
$$(M, \mathcal{G}, \mu) = (\Omega \times [t, T], \mathcal{F} \otimes \mathcal{B}([t, T]), \mathbb{P} \otimes \lambda)$$
for $t \in [0, T)$, or just $(M, \mathcal{G}, \mu) = (\Omega, \mathcal{F}, \mathbb{P})$.

- In the second part of the definition we implicitly require that $X(\omega, \cdot) \in L^1_{loc}(\mathbb{R})$ for almost every ω.

- Due to the definition of the weak derivative being an element of L^1_{loc}, we know that $\partial_x f$ and $\partial_x X(\omega, x)$ describe equivalence classes. Hence there are many versions for the weak derivative that satisfy equation (2.27). In some applications it is important which version is considered. In the case of stochastic processes, for instance, one would like to have a measurable version. This indeed exists if X is a measurable function of (ω, t, x), i.e. there exists a version of $\partial_x X(\omega, x)$ preserving this property (cf. [6, p. 18]).

- Weakly differentiable functions have a continuous version according to section 1.1.3 in [10], i.e. one can modify the function on a null set to obtain a continuous function. This version is even absolutely continuous and has obviously the same weak derivative. Nevertheless, this result is only true if the domain of the function is one-dimensional, but there is a weaker result for the multidimensional case.

The above definition motivates the next simple statement.

Proposition 2.31. *Let $A \subseteq \mathbb{R}$ be open and $f, g : A \to \mathbb{R}$ be weakly differentiable. Then $f + g$ is also weakly differentiable and $\partial_x(f + g) = \partial_x f + \partial_x g$.*

Proof. Let $\varphi \in C_c^\infty(A)$. Due to the weak differentiability of f and g we have

$$\begin{aligned}\int_A \varphi'(x)\left(f(x) + g(x)\right)\,\mathrm{d}x &= \int_A \varphi'(x) f(x)\,\mathrm{d}x + \int_A \varphi'(x) g(x)\,\mathrm{d}x \\ &= -\int_A \varphi(x)\partial_x f(x)\,\mathrm{d}x - \int_A \varphi(x)\partial_x g(x)\,\mathrm{d}x \\ &= -\int_A \varphi(x)\left(\partial_x f(x) + \partial_x g(x)\right)\,\mathrm{d}x.\end{aligned}$$

Therefore, $f + g$ is weakly differentiable with $\partial_x f + \partial_x g$ being a version of the weak derivative. Note that all the integrals above exist since $f, g \in L^1_{loc}(A)$. □

There is a version of the fundamental theorem of calculus for the weak setting. This statement and its proof can be found in Lemma A.2.1 of [6].

Proposition 2.32. *Let $(M, \mathcal{G}, \mu)$ be a finite measure space and let $X : M \times \mathbb{R} \to \mathbb{R}$ be a measurable function that is weakly differentiable w.r.t. the second component. Then for almost all $x_1, x_2 \in \mathbb{R}$ and almost all $\omega \in M$ we have*

$$\int_{x_1}^{x_2} \partial_x X(\omega, y)\,\mathrm{d}y = X(\omega, x_2) - X(\omega, x_1). \tag{2.28}$$

We now generalize Lemma A.2.3 of [6] for the purpose of stating Proposition 2.35, Proposition 2.36 and Proposition 2.37 that enable us to interchange weak differentiation and integration.

Lemma 2.33. *Let $(M, \mathcal{G}, \mu)$ be a finite measure space and let $X : M \times \mathbb{R} \to \mathbb{R}$ be a measurable function that is weakly differentiable w.r.t. the second component. Let $p \geq 1$. Furthermore, assume that*

(1) $\int_M |X(\cdot, x)|^p\,\mathrm{d}\mu < \infty$ *for all* $x \in \mathbb{R}$,

(2) $\operatorname*{ess\,sup}_{x \in K} \int_M |\partial_x X(\cdot, x)|^p\,\mathrm{d}\mu < \infty$ *for all compact sets* $K \subseteq \mathbb{R}$.

Then the mapping $x \mapsto \int_M |X(\cdot, x)|^p\,\mathrm{d}\mu$ is locally bounded, meaning that for all compact sets $K \subseteq \mathbb{R}$ we have

$$\operatorname*{ess\,sup}_{x \in K} \int_M |X(\cdot, x)|^p\,\mathrm{d}\mu < \infty.$$

Proof. Let $K \subseteq \mathbb{R}$ be an arbitrary compact set. We choose $x_0 \in \mathbb{R}$ such that equation (2.28) holds true for almost all $(\omega, x) \in M \times \mathbb{R}$ and $x_0 < \inf K$. Then we have for almost every $x \in K$ that

$$\begin{aligned}
&\frac{1}{2^{p-1}} \int_M |X(\omega, x)|^p \, \mathrm{d}\mu(\omega) \\
&\quad \leq \frac{1}{2^{p-1}} \int_M 2^{p-1} \left(\left(\int_{x_0}^x |\partial_x X(\omega, y)| \, \mathrm{d}y \right)^p + |X(\omega, x_0)|^p \right) \mathrm{d}\mu(\omega) \\
&\quad \leq \int_M \left((x - x_0)^{p-1} \int_{x_0}^x |\partial_x X(\omega, y)|^p \, \mathrm{d}y + |X(\omega, x_0)|^p \right) \mathrm{d}\mu(\omega) \\
&\quad = (x - x_0)^{p-1} \int_{x_0}^x \int_M |\partial_x X(\omega, y)|^p \, \mathrm{d}\mu(\omega) \, \mathrm{d}y + \int_M |X(\omega, x_0)|^p \, \mathrm{d}\mu(\omega) \\
&\quad \leq |x - x_0|^p \left(\operatorname*{ess\,sup}_{z \in K} \int_M |\partial_x X(\omega, z)|^p \, \mathrm{d}\mu \right) + \int_M |X(\omega, x_0)|^p \, \mathrm{d}\mu(\omega) \\
&\quad \leq \left(\max_{y \in K} |y - x_0|^p \right) \left(\operatorname*{ess\,sup}_{z \in K} \int_M |\partial_x X(\omega, z)|^p \, \mathrm{d}\mu(\omega) \right) + \int_M |X(\omega, x_0)|^p \, \mathrm{d}\mu(\omega) < \infty,
\end{aligned}$$

where we have used Proposition 2.32, two times Jensen's inequality, Fubini's theorem and assumptions (1) and (2). Finally, we obtain

$$\operatorname*{ess\,sup}_{x \in K} \int_M |X(\cdot, x)|^p \, \mathrm{d}\mu < \infty.$$

□

Remark 2.34. Note that Lemma 2.33 implies that the mapping $x \mapsto \int_M |X(\cdot, x)|^p \, \mathrm{d}\mu$ is locally integrable, i.e. $\int_M |X(\omega, \cdot)|^p \, \mathrm{d}\mu(\omega) \in L^1_{loc}(\mathbb{R})$. Thus, we can generalize Lemma A.2.4, Lemma A.2.5 and Lemma A.2.6 in [6], which admit the weak differentiation of the conditional expectation, the Lebesgue integral and the stochatic integral w.r.t. the Brownian motion. In particular, we only require in (3) of Proposition 2.35, Proposition 2.36 and Proposition 2.37 that the terms are essentially bounded on all compact sets instead of demanding global essential boundedness. The proofs of these statements are analogous to the ones presented in [6]. One just has to apply Lemma 2.33 instead of Lemma A.2.3. Consequently, we omit the proofs here.

The next three statements are generalizations of Lemma A.2.4, Lemma A.2.5 and Lemma A.2.6 in [6]. The proofs are analogous to the ones presented in [6] as mentioned in Remark 2.34.

Proposition 2.35. *Let $X : \Omega \times \mathbb{R} \to \mathbb{R}$ be measurable such that*

(1) X is weakly differentiable in x,

(2) $\mathbb{E}[|X(\cdot, x)|] < \infty$ for all $x \in \mathbb{R}$,

(3) $\operatorname*{ess\,sup}_{x \in K} \mathbb{E}[|\partial_x X(\cdot, x)|] < \infty$ for all compact sets $K \subseteq \mathbb{R}$.

Let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra. Then the mapping $(\omega, x) \mapsto \mathbb{E}[X(\cdot, x)|\mathcal{G}](\omega)$ is measurable, weakly differentiable w.r.t. x and $\partial_x \mathbb{E}[X(\cdot, x)|\mathcal{G}] = \mathbb{E}[\partial_x X(\cdot, x)|\mathcal{G}]$.

Proposition 2.36. *Let $Z : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ be measurable such that*

(1) Z is weakly differentiable w.r.t. x,

(2) $\mathbb{E}\left[\int_0^T |Z_s(\cdot, x)| \,\mathrm{d}s\right] < \infty$ for all $x \in \mathbb{R}$,

(3) $\operatorname*{ess\,sup}_{x\in K} \mathbb{E}\left[\int_0^T |\partial_x Z_s(\cdot, x)| \,\mathrm{d}s\right] < \infty$ for all compact sets $K \subseteq \mathbb{R}$.

Then the mapping $(\omega, x) \mapsto X(\omega, x) := \int_0^T Z_s(\omega, x) \,\mathrm{d}s$ is measurable, weakly differentiable w.r.t. x and $\partial_x X(\omega, x) = \int_0^T \partial_x Z_s(\omega, x) \,\mathrm{d}s$.

Proposition 2.37. *Let $Z : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ be progressively measurable such that*

(1) Z is weakly differentiable w.r.t. x,

(2) $\mathbb{E}\left[\int_0^T |Z_s(\cdot, x)|^2 \,\mathrm{d}s\right] < \infty$ for all $x \in \mathbb{R}$,

(3) $\operatorname*{ess\,sup}_{x\in K} \mathbb{E}\left[\int_0^T |\partial_x Z_s(\cdot, x)|^2 \,\mathrm{d}s\right] < \infty$ for all compact sets $K \subseteq \mathbb{R}$.

Then the mapping $(\omega, x) \mapsto X(\omega, x) := \int_0^T Z_s(\omega, x) \,\mathrm{d}W_s(\omega)$ is measurable, weakly differentiable w.r.t. x and $\partial_x X(\omega, x) = \int_0^T \partial_x Z_s(\omega, x) \,\mathrm{d}W_s(\omega)$.

Depending on the particular framework, there are several chain rules for weak derivatives. Here we present four different ones.

Proposition 2.38. *Let $n \in \mathbb{N}$ and $g : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz continuous function. Furthermore, let $f_i : \mathbb{R} \to \mathbb{R}$ be a weakly differentiable function for $i = 1, \ldots, n$. Then the composite function $g(f_1(\cdot), \ldots, f_n(\cdot))$ is also weakly differentiable. Moreover, if g is everywhere classically differentiable we have*

$$\partial_x \left(g(f_1(x), \ldots, f_n(x))\right) = \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f_1(x), \ldots, f_n(x))\partial_x f_i(x) \tag{2.29}$$

for almost every $x \in \mathbb{R}$.

The proof of this statement is presented in Lemma A.3.1 of [6]. Note that the statement still holds true if g is only locally Lipschitz continuous and the functions $f_1, \ldots, f_n$ are bounded. Moreover, it is valid if g is not globally Lipschitz continuous but continuously differentiable.

Corollary 2.39. *For $n \in \mathbb{N}$ let $g : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function and $f_i : \mathbb{R} \to \mathbb{R}$ be a weakly differentiable function for $i = 1, \ldots, n$. Then the composite function $g(f_1(\cdot), \ldots, f_n(\cdot)) : \mathbb{R} \to \mathbb{R}$ is weakly differentiable and*

$$\partial_x \left(g(f_1(x), \ldots, f_n(x))\right) = \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f_1(x), \ldots, f_n(x))\partial_x f_i(x)$$

for almost every $x \in \mathbb{R}$.

Proof. We have to show that (2.27) holds. Let therefore $\varphi \in C_c^\infty(\mathbb{R})$ and let $I_1 \subseteq \mathbb{R}$ be a non-empty compact interval such that $\varphi(x) = 0$ for all $\mathbb{R} \setminus I_1$. Moreover, let $I_2 \subseteq \mathbb{R}$ be an open, bounded interval such that $I_1 \subseteq I_2$. Note that:

- According to the definition of weak differentiability we have $f_i, \partial_x f_i \in L^1_{loc}(\mathbb{R})$ for all $i = 1, \ldots, n$. In particular, we obtain that $f_i, \partial_x f_i \in L^1(I_2)$.
- Theorem 2.1.4 in [14] implies that on I_2 there are representatives of $f_1, \ldots, f_n$ that are absolutely continuous. Without loss of generality we assume that the functions $f_1, \ldots, f_n$ restricted to I_2 are already absolutely continuous, since these representatives are anyway equal to those function almost everywhere. Consequently, $f_1, \ldots, f_n$ are also continuous on $I_1 \subseteq I_2$.
- The set $f_i(I_1)$ is a compact interval by the continuity of the functions f_i for all $i = 1, \ldots, n$. Thus, we can define the compact set $K := f_1(I_1) \times \ldots \times f_n(I_1)$.
- $g|_K$ is Lipschitz continuous because g is continuously differentiable on $\mathbb{R}^n$.

Now we observe that

$$\begin{aligned}\int_{\mathbb{R}} \varphi'(x) g(f_1(x), \ldots, f_n(x)) \, \mathrm{d}x &= \int_{I_1} \varphi'(x) g(f_1(x), \ldots, f_n(x)) \, \mathrm{d}x \\ &= -\int_{I_1} \varphi(x) \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f_1(x), \ldots, f_n(x)) \partial_x f_i(x) \, \mathrm{d}x \\ &= -\int_{\mathbb{R}} \varphi(x) \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f_1(x), \ldots, f_n(x)) \partial_x f_i(x) \, \mathrm{d}x.\end{aligned}$$

The first and third equality hold true since φ and φ' vanish outside of I_1. Moreover, the second one is valid by Proposition 2.38, because $f_1(I_1) \times \ldots \times f_n(I_1) = K$ and the function g restricted to K is Lipschitz continuous. We emphasize that the weak derivative of $g(f_1(\cdot), \ldots, f_n(\cdot))$ is indeed locally integrable on $\mathbb{R}$ since all the partial derivatives of g are bounded on all compact sets. Hence we have for all compact $A \subseteq \mathbb{R}$ and
$K' := f_1(A) \times \ldots \times f_n(A)$ that

$$\int_A \left| \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f_1(x), \ldots, f_n(x)) \partial_x f_i(x) \right| \, \mathrm{d}x \leq \sum_{i=1}^n \sup_{y \in K'} \left| \frac{\partial g}{\partial y_i}(y) \right| \int_A |\partial_x f_i(x)| \, \mathrm{d}x < \infty,$$

because by assumption f_i is weakly differentiable for all $i = 1, \ldots, n$. Finally, the arbitrary choice of φ provides the result. $\square$

If one interchanges the functions f and g in Proposition 2.38, the composite function $f \circ g$ stays weakly differentiable under additional conditions. That is a consequence of Theorem 2.2.2 in [14] on which the next statement is based.

Proposition 2.40. *Let $f : \mathbb{R} \to \mathbb{R}$ be weakly differentiable and $g : \mathbb{R} \to \mathbb{R}$ be Lipschitz continuous such that the inverse g^{-1} exists and is also Lipschitz continuous. Then $f \circ g$ is weakly differentiable and*

$$\partial_x \left[f(g(x)) \right] = \left(\partial_x f \right) (g(x)) g'(x),$$

for almost every $x \in \mathbb{R}$.

The last chain rule that we present here gives a representation of (2.29) when the function g is not everywhere differentiable. It is proven in Lemma A.3.2 of [6].

Proposition 2.41. *Let $(M, \mathcal{G}, \mu)$ be a finite measure space, $g : M \times \mathbb{R} \to \mathbb{R}$ be measurable and Lipschitz continuous in the second component with Lipschitz constant $L_{g,x}$. Moreover, let $X : M \times \mathbb{R} \to \mathbb{R}$ be measurable and weakly differentiable w.r.t. the second variable x. Then also $g \circ X$ is weakly differentiable w.r.t. x, i.e. for almost all $\omega \in M$ the mapping $g(\omega, X(\omega, \cdot)) : \mathbb{R} \to \mathbb{R}$ is weakly differentiable. Furthermore, there exists a mapping $\Delta g : M \times \mathbb{R} \to \mathbb{R}$ such that $|\Delta g| \leq L_{g,x}$ and*

$$\partial_x \left(g(\omega, X(\omega, x)) \right) = \Delta g(\omega, x) \partial_x X(\omega, x)$$

for almost all $(\omega, x) \in M \times \mathbb{R}$.

We conclude this chapter with a statement that connects weak differentiability with the classical one.

Proposition 2.42. *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and weakly differentiable and let furthermore g be a version of the weak derivative that is continuous. Then f is continuously differentiable and $f' = g$.*

Since weakly differentiable functions have a continuous version as stated in Remark 2.30, we actually do not have to require continuity of f. However, we want to emphasize that we consider this continuous version, which will be continuously differentiable.

In order to prove this proposition we present a useful fact about the approximation of weak derivatives. Therefore, we define for $\varepsilon > 0$ and $x \in \mathbb{R}$ the function $\varphi_\varepsilon(x) := \frac{1}{\varepsilon}\varphi(\frac{x}{\varepsilon})$, where φ is given by

$$\varphi(x) := \begin{cases} C \exp\left(-\frac{1}{1-|x|^2}\right) & \text{, if } |x| < 1 \\ 0 & \text{, else} \end{cases}$$

and the constant C is chosen such that $\int_{\mathbb{R}} \varphi(x)\, \mathrm{d}x = 1$. For a function $f \in L^1_{loc}(\mathbb{R})$ the so-called *regularizer* f_ε is defined by

$$f_\varepsilon(x) := (\varphi_\varepsilon * f)(x) = \int_{\mathbb{R}} \varphi_\varepsilon(x-y) f(y)\, \mathrm{d}y,\ x \in \mathbb{R}.$$

Lemma 2.43. *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Then for all $\varepsilon > 0$ we have $f_\varepsilon \in C^\infty(\mathbb{R})$ and $f'_\varepsilon = \varphi'_\varepsilon * f$. Furthermore, the regularizer f_ε converges uniformly to f on each compact subset of $\mathbb{R}$ as $\varepsilon \to 0$.*

The proof of this lemma can be found in Theorem 1.6.1 of [14]. We turn to the proof of Proposition 2.42.

Proof of Proposition 2.42. It suffices to show that f is continuously differentiable on every compact subset of $\mathbb{R}$. Let $K \subseteq \mathbb{R}$ be compact. According to Lemma 2.43, the regularizer f_ε converges uniformly to f on K as $\varepsilon \to 0$. Moreover, for all $\varepsilon > 0$ we have $f_\varepsilon \in C^1(K)$ and

$$\begin{aligned} f'_\varepsilon(x) &= (\varphi'_\varepsilon * f)(x) = \int_{\mathbb{R}} \varphi'_\varepsilon(x-y) f(y)\, \mathrm{d}y \\ &= \int_{\mathbb{R}} \varphi_\varepsilon(x-y) g(y)\, \mathrm{d}y = (\varphi_\varepsilon * g)(x) = g_\varepsilon(x),\ x \in K, \end{aligned}$$

where we have used that g is a weak derivative of f. Applying Lemma 2.43 to the function g yields that the regularizer $g_\varepsilon = f'_\varepsilon$ converges uniformly to g on K as $\varepsilon \to 0$. Altogether, the uniform convergence of f_ε and f'_ε to f and g, respectively, implies that the sequence $(\phi_n)_{n\in\mathbb{N}}$, defined by $\phi_n := f_{\frac{1}{n}}$, is a Cauchy sequence in the space $C^1(K)$ equipped with the norm

$$\|h\| := \|h\|_\infty + \|h'\|_\infty, \; h \in C^1(K).$$

This space is complete and thus $(\phi_n)_{n\in\mathbb{N}}$ converges to a limit $\tilde{f} \in C^1(K)$. But $(\phi_n)_{n\in\mathbb{N}}$ also converges uniformly to f by the considerations above. Therefore, the uniqueness of the limit implies $f = \tilde{f}$. In the same way the sequence $(\phi'_n)_{n\in\mathbb{N}}$ converges uniformly to g and this implies $\tilde{f}' = f' = g$. To sum up, we observe that f is continuously differentiable on $\mathbb{R}$ and $f' = g$. $\square$

Chapter 3

Optimal control of diffusion coefficients in a bounded setting

In this chapter we study the optimal control problem presented in [1] with some small changes. Unlike this publication we assume that the controls are bounded and that the function f, characterizing the running costs, does not depend on the state of our underlying process M. These requirements cause various changes in the course of action and thus the statements are slightly different. Nevertheless, our setting allows us to follow mainly the same train of thought.

We consider a stochastic process M with dynamics given by

$$M_t^\alpha = \mu(t, M_t^\alpha)dt + \alpha_t dW_t,$$

where the drift coefficient $\mu : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ is an affine linear function in the last variable and the diffusion is directly controlled by a suitable control α. In the optimal control problem, which we study here, we aim at minimizing the target function

$$\mathbb{E}\left[\int_0^T f(s, \alpha_s)\,\mathrm{d}s + g(M_T^\alpha)\right]$$

over all appropriate controls α, which take values in a given compact interval A only. The functions f and g are especially assumed to be convex in α and M, respectively, and are allowed to depend on ω.

One can think of M as the position of a particle in a medium with temperature α. Any temperature change influences the fluctuation of the particle. By heating or cooling the medium one can try to steer the particle into a target area. In our bounded setting, however, one cannot arbitrarily increase or decrease the medium's temperature since only temperatures in a fixed compact interval are possible. That seems reasonable in practice. Every temperature change involves costs characterized by the function f. If a control does not steer the particle into the right position at time T, the function g can be understood as penalization. We are aiming to minimize the costs, while steering the particle as good as possible.

We minimize the target function by determining an optimal control with the help of Pontryagin's maximum principle. Therefore, we have to solve the adjoint FBSDE that is coupled and unfortunately does not satisfy standard Lipschitz conditions (SLC). Nevertheless, we can consider a suitable auxiliary FBSDE that fulfils SLC, and consequently, we can apply the method of decoupling fields. This technique enables us to prove solvability of the auxiliary FBSDE on the whole time interval $[0,T]$ and involves the following steps. Firstly, we assume on the contrary that there exists only a solution to the auxiliary FBSDE on an interval $I_{max} = (t_{min}, T]$. We calculate a uniform bound for the Lipschitz constant of the decoupling field using the gradient process. This contradicts our assumption according to Proposition 2.22. Hence there exists a unique solution to the auxiliary FBSDE that we transfer to our original one. We emphasize that allowing only bounded controls makes this transformation more complicated.

3.1 Introducing the problem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and let W be a Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $T > 0$ be a finite time horizon. Define furthermore the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ by $\mathcal{F}_t := \sigma(\mathcal{N}, (W_s)_{s\in[0,t]})$ for $t \in [0,T]$ with $\mathcal{N}$ being the set of all $\mathbb{P}$-null sets in $\mathcal{F}$.

Let $A := [l,r] \subseteq \mathbb{R}$ be a compact, non-empty interval that represents all possible control values. We define $\mathcal{A}$ as the set of all progressively measurable processes $\alpha : \Omega \times [0,T] \to A$ with

$$\mathbb{E}\left[\int_0^T \alpha_s^2 \, \mathrm{d}s\right] < \infty.$$

We call a control α admissible if $\alpha \in \mathcal{A}$. Moreover, let $b, B : \Omega\times[0,T] \to \mathbb{R}$ be progressively measurable and bounded processes. For the initial value $m \in \mathbb{R}$ and the control $\alpha \in \mathcal{A}$ we consider the controlled diffusion process $M^{m,\alpha}$ driven by the stochastic differential equation

$$M_t^{m,\alpha} = m + \int_0^t (b_s + B_s M_s^{m,\alpha}) \, \mathrm{d}s + \int_0^t \alpha_s \, \mathrm{d}W_s. \tag{3.1}$$

Note that our assumptions on the processes $(b_t)_{t\in[0,T]}, (B_t)_{t\in[0,T]}$ and on the control processes provide the existence and uniqueness of a strong solution to (3.1) starting from m at time 0 (see e.g. Theorem 1.3.15 in [11]). We denote this solution by $M^{m,\alpha} = (M_t^{m,\alpha})_{t\in[0,T]}$.

Let $g : \Omega \times \mathbb{R} \to \mathbb{R}$ and $f : \Omega \times [0,T] \times A \to \mathbb{R}$ be measurable functions. We require that the mapping $(\omega, t) \mapsto f(\omega, t, a)$ is progressively measurable for all $a \in A$. Moreover, we assume that f and g satisfy the following:

(C1) For all $(\omega, t) \in \Omega \times [0,T]$ the mappings $f(\omega, t, \cdot)$ and $g(\omega, \cdot)$ are strictly convex. Moreover,

$$\|f(\cdot,\cdot,0)\|_\infty + \|g(\cdot,0)\|_\infty < \infty,$$

where $\|h\|_\infty$ denotes the essential supremum of $|h|$ for some measurable function h.

(C2) The functions $g(\omega, \cdot)$ and $f(\omega, t, \cdot)$ are two times continuously differentiable for all $(\omega, t) \in \Omega \times [0,T]$. Additionally, the derivatives g' and f_a are Lipschitz continuous in the variables m and a, respectively, i.e. there exists a constant $L \geq 0$ such that for all $\omega \in \Omega$, $t \in [0,T]$, $m_1, m_2 \in \mathbb{R}$, $a_1, a_2 \in A$

$$|g'(\omega, m_1) - g'(\omega, m_2)| \leq L|m_1 - m_2|,$$
$$|f_a(\omega, t, a_1) - f_a(\omega, t, a_2)| \leq L|a_1 - a_2|.$$

Furthermore, it holds

$$\|f_a(\cdot,\cdot,0)\|_\infty + \|g'(\cdot,0)\|_\infty < \infty.$$

(C3) There is a constant $\delta_l > 0$ such that

$$f_{aa}(\omega,t,a) \geq \delta_l \qquad \text{and} \qquad g''(\omega,m) \geq \delta_l$$

for all $(\omega,t,a,m) \in \Omega \times [0,T] \times A \times \mathbb{R}$.

Remark 3.1. Note that we often omit the ω-argument of the functions f and g to simplify the notation. Furthermore, condition (C2) implies that f_{aa} and g'' are bounded from above, because for all $(\omega,t,a,m) \in \Omega \times [0,T] \times A \times \mathbb{R}$ one has

$$0 \leq f_{aa}(\omega,t,a) = \lim_{h\to 0} \frac{1}{|h|} |f_a(\omega,t,a+h) - f_a(\omega,t,a)| \leq \lim_{h\to 0} \frac{1}{|h|} L|h| = L,$$

and

$$0 \leq g''(\omega,m) = \lim_{h\to 0} \frac{1}{|h|} |g'(\omega,m+h) - g'(\omega,m)| \leq L.$$

We denote by $\delta_u \in [\delta_l,\infty)$ a constant that bounds both f_{aa} and g'', i.e.

$$\max\left(f_{aa}(\omega,t,a), g''(\omega,m)\right) \leq \delta_u,$$

for all $(\omega,t,a,m) \in \Omega \times [0,T] \times A \times \mathbb{R}$.

One can think of the function f as the running costs, whereas the function g can be interpreted as the terminal costs of our minimization problem (P). The target function J describes our total costs and is given by

$$J(m,\alpha) := \mathbb{E}\left[\int_0^T f\left(s,\alpha_s\right) \,\mathrm{d}s + g\left(M_T^{m,\alpha}\right)\right]$$

for $m \in \mathbb{R}$ and $\alpha \in \mathcal{A}$. This function J is well-defined because one can show as in Remark 2.24 that the expected value exists using (C1) and (C2). Our goal is now to solve the following problem.

Problem (P): Minimize the target function J over all admissible controls, i.e. determine

$$\inf_{\alpha\in\mathcal{A}} J(m,\alpha),$$

and find an optimal control $\hat{\alpha} \in \mathcal{A}$, i.e. a control $\hat{\alpha}$ satisfying $J(m,\hat{\alpha}) = \inf_{\alpha\in\mathcal{A}} J(m,\alpha)$.

3.2 Deriving the FBSDE via the maximum principle

We aim at solving (P) with the help of the Pontryagin's maximum principle, which is presented in section 2.4. To that end, we introduce the *Hamiltonian* $\mathcal{H}$, defined by

$$\mathcal{H}(\omega, t, m, a, y, z) := (b_t(\omega) + B_t(\omega)m)y + az + f(\omega, t, a),$$

for $(\omega, t, m, a, y, z) \in \Omega \times [0, T] \times \mathbb{R} \times A \times \mathbb{R} \times \mathbb{R}$. We notice that $\mathcal{H}(\omega, t, \cdot, \cdot, y, z)$ is two times continuously differentiable and convex for all $\omega \in \Omega, t \in [0, T], y, z \in \mathbb{R}$, due to the properties of the function f. The Hamiltonian is even strictly convex in the variable a. Now we consider the possibly coupled adjoint FBSDE given by

$$\begin{aligned} M_t^{m,\alpha} &= m + \int_0^t (b_s + B_s M_s^{m,\alpha}) \,\mathrm{d}s + \int_0^t \alpha_s \,\mathrm{d}W_s, \\ Y_t^{m,\alpha} &= g'(M_T^{m,\alpha}) + \int_t^T B_s Y_s^{m,\alpha} \,\mathrm{d}s - \int_t^T Z_s^{m,\alpha} \,\mathrm{d}W_s, \ t \in [0, T], \end{aligned} \tag{3.2}$$

where $\alpha \in \mathcal{A}$. The solvability of this equation essentially depends on the control α. However, we are only interested in a solution if α is a candidate for an optimal control, because we want to apply the maximum principle. To determine such a control we calculate the minimum of $\mathcal{H}$ over all $a \in A$.

Lemma 3.2. *The Hamiltonian $\mathcal{H}$ is minimized by $a_*(t, -z)$, i.e. we have*

$$\min_{a \in A} \mathcal{H}(t, m, a, y, z) = \mathcal{H}(t, m, a_*(t, -z), y, z)$$

for all $(\omega, t, m, z) \in \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}$, where $a_* : \Omega \times [0, T] \times \mathbb{R} \to \mathbb{R}$ is given by

$$a_*(t, z) := \begin{cases} l & , z < f_a(t, l), \\ f_a^{-1}(t, z) & , z \in [f_a(t, l), f_a(t, r)] \,, \\ r & , z > f_a(t, r). \end{cases}$$

The function f_a^{-1} denotes the inverse of the partial derivative f_a w.r.t. the last component. This inverse indeed exists since f_a is strictly increasing by (C3).

Proof. Let $(\omega, t, m, z) \in \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}$ be fixed. We consider now $\mathcal{H}(t, m, a, y, z)$ as function in a and determine its minimum depending on (ω, t, m, z). From assumption (C2) and the definition of the Hamiltonian we know that $\mathcal{H}$ is two times continuously differentiable in a with derivatives

$$\begin{aligned} \mathcal{H}_a(t, m, a, y, z) &= z + f_a(t, a) \\ \mathcal{H}_{aa}(t, m, a, y, z) &= f_{aa}(t, a) \end{aligned}$$

for $a \in A$. Consequently, $\mathcal{H}$ is strictly convex with $\mathcal{H}_{aa} \geq \delta_l$ and the derivative $\mathcal{H}_a$ is strictly increasing according to (C1) and (C3). In the same way we see by (C1)-(C3) that the function f_a is strictly increasing in a and has the image $[f_a(t, l), f_a(t, r)]$.

We distinguish three cases:

1. If $z < -f_a(t,r)$, then we obtain

$$\mathcal{H}_a(t,m,a,y,z) = z + f_a(t,a) \leq z + f_a(t,r) < 0$$

   for all $a \in A$, since f_a is strictly increasing. Consequently, the Hamiltonian $\mathcal{H}$ is strictly decreasing in a and thus attains its minimum at $a = r$.

2. If $z \in [-f_a(t,r), -f_a(t,l)]$, then there exists an $a_0 \in A$ such that $z = -f_a(t,a_0)$ and therefore

$$\mathcal{H}_a(t,m,a_0,y,z) = z + f_a(t,a_0) = 0.$$

   Hence the Hamiltonian $\mathcal{H}$ attains its minimum at $a = a_0$ because of the strict convexity. We observe that $a_0 = f_a^{-1}(t,-z)$, since f_a is a bijection between A and $[f_a(t,l), f_a(t,r)]$.

3. If $z > -f_a(t,l)$, then it holds that

$$\mathcal{H}_a(t,m,a,y,z) = z + f_a(t,a) \geq z + f_a(t,l) > 0$$

   for all $a \in A$, since f_a is strictly increasing. Therefore, the Hamiltonian $\mathcal{H}$ is also strictly increasing in a and thus attains its minimum at $a = l$.

Summarizing above result we see that

$$\min_{a\in A} \mathcal{H}(t,m,a,y,z) = \begin{cases} \mathcal{H}(t,m,r,y,z) & ,z < -f_a(t,r), \\ \mathcal{H}(t,m,f_a^{-1}(t,-z),y,z) & ,z \in [-f_a(t,r), -f_a(t,l)]\,, \\ \mathcal{H}(t,m,l,y,z) & ,z > -f_a(t,l). \end{cases}$$

The minimum does not depend on m but on (ω, t, z). □

Now that we have found the function a_* that minimizes the Hamiltonian $\mathcal{H}$, we study its measurability properties.

Lemma 3.3. *The mapping a_*, defined in Lemma 3.2, is measurable and for all $z \in \mathbb{R}$ we have that $a_*(\cdot,\cdot,z)$ is progressively measurable. Moreover, for any progressively measurable process Z the composition $a_*(\cdot,\cdot,Z_\cdot(\cdot))$ is progressively measurable.*

Proof. In order to prove the measurability of the function a_*, we fix $a \in \mathbb{R}$. Note that

$$\begin{aligned} &\{(\omega,s,z) \in \Omega \times [0,T] \times \mathbb{R} : a_*(\omega,s,z) \leq a\} \\ &\quad = \left\{ \begin{array}{ll} \emptyset & ,a < l \\ \Omega \times [0,T] \times \mathbb{R} & ,a \geq r \end{array} \right\} \in \mathcal{F} \otimes \mathcal{B}([0,T]) \otimes \mathcal{B}(\mathbb{R}), \end{aligned}$$

because a_* takes values in $A = [l,r]$ only. Considering now the case $a \in [l,r)$ yields

$$\begin{aligned} &\{(\omega,s,z) \in \Omega \times [0,T] \times \mathbb{R} : a_*(\omega,s,z) \leq a\} \\ &= \{(\omega,s,z) : z < f_a(\omega,s,l)\} \cup \{(\omega,s,z) : f_a^{-1}(\omega,s,z) \leq a, z \in [f_a(\omega,s,l), f_a(\omega,s,r)]\} \\ &= \{(\omega,s,z) : z \leq f_a(\omega,s,a)\} \in \mathcal{F} \otimes \mathcal{B}([0,T]) \otimes \mathcal{B}(\mathbb{R}). \end{aligned}$$

We have used, in particular, that f_a is measurable and strictly increasing in a due to (C3). This shows the measurability of a_*.

Now we fix $z \in \mathbb{R}, t \in [0,T]$ and show that $(\omega,s) \mapsto a_*(\omega,s,z)$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t])$-measurable. Let $a \in \mathbb{R}$ and note that

$$\{(\omega,s) \in \Omega \times [0,t] : a_*(\omega,s,z) \le a\} = \left\{ \begin{array}{ll} \emptyset & , a < l \\ \Omega \times [0,t] & , a \ge r \end{array} \right\} \in \mathcal{F}_t \otimes \mathcal{B}([0,t]).$$

Furthermore, if $a \in [l,r)$ we have

$$\begin{aligned} &\{(\omega,s) \in \Omega \times [0,t] : a_*(\omega,s,z) \le a\} \\ &\quad = \{(\omega,s) \in \Omega \times [0,t] : z < f_a(\omega,s,l)\} \\ &\qquad \cup \{(\omega,s) \in \Omega \times [0,t] : f_a^{-1}(\omega,s,z) \le a, z \in [f_a(\omega,s,l), f_a(\omega,s,r)]\} \\ &\quad = \{(\omega,s) \in \Omega \times [0,t] : z \le f_a(\omega,s,a)\} \in \mathcal{F}_t \otimes \mathcal{B}([0,t]). \end{aligned}$$

Here we have used the progressive measurability of f_a and again the strict monotonicity of f_a in a.

Let now Z be a progressively measurable process and $t \in [0,T]$. We show that $a_*(\cdot,\cdot,Z_\cdot(\cdot))$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t])$-measurable. Note that $a_*(\omega,s,\cdot)$ is continuous for all fixed $(\omega,s) \in \Omega \times [0,T]$. Hence we can approximate a_* pointwise by the sequence

$$a_*^{(n)}(\omega,s,z) := \sum_{j \in \mathbb{Z}} \mathbb{1}_{\left(\frac{j-1}{n},\frac{j}{n}\right]}(z) a_*\left(\omega,s,\tfrac{j}{n}\right), \ n \in \mathbb{N}, \ (\omega,s,z) \in \Omega \times [0,T] \times \mathbb{R}.$$

We show that $a_*^{(n)}(\cdot,\cdot,Z_\cdot(\cdot))$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t])$-measurable: For all $n \in \mathbb{N}$ and $a \in \mathbb{R}$ we have

$$\begin{aligned} &\{(\omega,s) \in \Omega \times [0,t] : a_*^{(n)}(\omega,s,Z_s(\omega)) \le a\} \\ &\quad = \bigcup_{j \in \mathbb{Z}} \left\{(\omega,s) \in \Omega \times [0,t] : a_*\left(\omega,s,\tfrac{j}{n}\right) \le a, Z_s(\omega) \in \left(\tfrac{j-1}{n},\tfrac{j}{n}\right]\right\} \in \mathcal{F}_t \otimes \mathcal{B}([0,t]), \end{aligned}$$

because Z and $a_*\left(\omega,s,\frac{j}{n}\right)$ are progressively measurable for all $j \in \mathbb{Z}, n \in \mathbb{N}$. Consequently, also the mapping $\Omega \times [0,t] \to \mathbb{R}, (\omega,s) \mapsto a_*(\omega,a,Z_s(\omega))$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t])$-measurable as a pointwise limit of the measurable functions $a_*^{(n)}$. Finally, $a_*(\cdot,\cdot,Z_\cdot(\cdot))$ is progressively measurable. □

Lemma 3.2 suggests that $\hat\alpha := (a_*(t,-Z_t^m))_{t \in [0,T]}$ is an optimal control for (P) if the triple (M^m, Y^m, Z^m) solves the FBSDE

$$\begin{aligned} M_t^m &= m + \int_0^t (b_s + B_s M_s^m)\,\mathrm{d}s + \int_0^t a_*(s,-Z_s^m)\,\mathrm{d}W_s, \\ Y_t^m &= g'(M_T^m) + \int_t^T B_s Y_s^m\,\mathrm{d}s - \int_t^T Z_s^m\,\mathrm{d}W_s, \ \ s \in [0,T]. \end{aligned} \tag{3.3}$$

The next proposition states that this is indeed the case. Note that (3.3) is the FBSDE (3.2) controlled by $\hat\alpha$, but in the following we omit the dependence on the control in the superscript of (M^m, Y^m, Z^m).

Proposition 3.4. *Suppose that there exists a solution (M^m, Y^m, Z^m) to (3.3) starting in $m \in \mathbb{R}$. Then $\hat\alpha \in \mathcal{A}$ is an optimal control for (P), where $\hat\alpha$ is given by*

$$\hat\alpha_t := a_*(t,-Z_t^m), \ t \in [0,T].$$

Proof. First of all note that $\hat{\alpha}$ is progressively measurable as a composition of the progressively measurable mappings a_* and Z^m (see Lemma 3.3). Moreover, we observe that $\hat{\alpha}_s(\omega) \in A$ for all $(\omega, s) \in \Omega \times [0, T]$, which implies

$$\mathbb{E}\left[\int_0^T \hat{\alpha}_s^2 \, \mathrm{d}s\right] \leq T(l^2 \vee r^2) < \infty.$$

This means that we have $\hat{\alpha} \in \mathcal{A}$.

In order to prove the optimality of $\hat{\alpha}$, we basically sum up the results mentioned above. For all $\omega \in \Omega, t \in [0, T]$ we observe that $x \mapsto \mathcal{H}(t, x, a, Y_t^m, Z_t^m)$ is convex for all $a \in A$, and

$$\mathcal{H}(t, M_t^m, \hat{\alpha}_t, Y_t^m, Z_t^m) = \min_{a \in A} \mathcal{H}(t, M_t^m, a, Y_t^m, Z_t^m)$$

by Lemma 3.2. Finally, the maximum principle (Theorem 2.25) yields that $\hat{\alpha}$ is optimal. $\square$

3.3 Transforming the FBSDE

We have seen in Proposition 3.4 that there exists an optimal control if we have a solution to the FBSDE (3.3). However, the question of the existence of such a global solution remains open. In order to prove that there is a solution on the whole interval $[0, T]$, we want to apply the method of decoupling field, which is introduced in section 2.3. Unfortunately, the parameters of (3.3) do not have to satisfy standard Lipschitz conditions (SLC), because there is no requirement for g' to fulfil $L_{g',x} < L_{\sigma,z}^{-1}$. But we can avoid this issue in considering the auxiliary FBSDE

$$\begin{aligned} X_t^x &= x + \int_0^t (b_s + B_s P_s^x) \, \mathrm{d}s + \int_0^t \left(\tilde{Z}_s^x - 2\gamma \tilde{a}_*^{-1}\left(s, \tilde{Z}_s^x\right)\right) \mathrm{d}W_s, \\ P_t^x &= \xi(X_T^x) - \int_t^T (b_s + B_s X_s^x) \, \mathrm{d}s - \int_t^T \tilde{Z}_s^x \, \mathrm{d}W_s, \ t \in [0, T], \end{aligned} \tag{3.4}$$

for $x \in \mathbb{R}$. Our goal is to show that this FBSDE satisfies SLC and has a solution. Moreover, we prove that one can recover a solution of our original FBSDE (3.3).

To understand the FBSDE (3.4) we define the following. Let $\gamma := \frac{1}{2\delta_u}$ and let the terminal condition $\xi : \Omega \times \mathbb{R} \to \mathbb{R}$ be given by

$$\begin{aligned} \xi(\omega, x) &:= \left((\mathrm{Id} - \gamma g'(\omega, \cdot)) \circ (\mathrm{Id} + \gamma g'(\omega, \cdot))^{-1}\right)(x) \\ &= (\mathrm{Id} + \gamma g'(\omega, \cdot))^{-1}(x) - \gamma g'\left(\omega, (\mathrm{Id} + \gamma g'(\omega, \cdot))^{-1}(x)\right) \\ &= 2(\mathrm{Id} + \gamma g'(\omega, \cdot))^{-1}(x) - (\mathrm{Id} + \gamma g'(\omega, \cdot))\left((\mathrm{Id} + \gamma g'(\omega, \cdot))^{-1}(x)\right) \\ &= 2(\mathrm{Id} + \gamma g'(\omega, \cdot))^{-1}(x) - x, \ (\omega, x) \in \Omega \times \mathbb{R}. \end{aligned} \tag{3.5}$$

Note that for all $\omega \in \Omega$ the functions $(\mathrm{Id} - \gamma g'(\omega, \cdot))$ and $(\mathrm{Id} + \gamma g'(\omega, \cdot))$ are continuously differentiable and the latter is even invertible since $g'' \geq 0$ and thus

$$(\mathrm{Id} + \gamma g'(\omega, \cdot))'(x) = 1 + \gamma g''(x) \geq 1, \ x \in \mathbb{R}.$$

Hence ξ is well-defined and continuously differentiable for all $\omega \in \Omega$. In addition, we define the function $\tilde{a}_* : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ by

$$\tilde{a}_*(\omega,t,z) := a_*(\omega,t,z) + \gamma z = \begin{cases} l + \gamma z & , z < f_a(\omega,t,l), \\ f_a^{-1}(\omega,t,z) + \gamma z & , z \in [f_a(\omega,t,l), f_a(\omega,t,r)], \\ r + \gamma z & , z > f_a(\omega,t,r). \end{cases}$$

Note that $\tilde{a}_*$ is invertible in z unlike the function a_*, because the mapping $z \mapsto f_a^{-1}(\omega,t,z)$ is strictly increasing for all $\omega \in \Omega,\ t \in [0,T]$ due to (C3). The inverse is given by the function $\tilde{a}_*^{-1} : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$, where

$$\begin{aligned} \tilde{a}_*^{-1}(\omega,t,z) &= \left(a_*(\omega,t,\cdot) + \gamma \mathrm{Id}\right)^{-1}(z) \\ &= \begin{cases} \frac{1}{\gamma}(z-l) & , z < l + \gamma f_a(\omega,t,l), \\ \left(f_a^{-1}(\omega,t,\cdot) + \gamma \mathrm{Id}\right)^{-1}(z) & , z \in [l + \gamma f_a(\omega,t,l), r + \gamma f_a(\omega,t,r)], \\ \frac{1}{\gamma}(z-r) & , z > r + \gamma f_a(\omega,t,r). \end{cases} \end{aligned}$$

This function has the following properties.

Lemma 3.5. *The function $\tilde{a}_*^{-1}$ is strictly increasing, measurable and for all $z \in \mathbb{R}$ the mapping $\tilde{a}_*^{-1}(\cdot,\cdot,z)$ is progressively measurable. Moreover, $\tilde{a}_*^{-1}$ is Lipschitz continuous in z with Lipschitz constant $\frac{1}{\gamma}$ and for all $z_1, z_2 \in \mathbb{R}, \omega \in \Omega, t \in [0,T]$ we have*

$$|\tilde{a}_*^{-1}(t,z_1) - \tilde{a}_*^{-1}(t,z_2)| \geq \frac{1}{\gamma + \frac{1}{\delta_l}} |z_1 - z_2|. \tag{3.6}$$

Proof. The mapping $\tilde{a}_*^{-1}(\omega,t,\cdot)$ is strictly increasing because $\tilde{a}_*(\omega,t,\cdot)$ is strictly increasing for all $(\omega,t) \in \Omega \times [0,T]$. Furthermore, for all $a \in \mathbb{R}$ we observe

$$\begin{aligned} &\{(\omega,t,z) \in \Omega \times [0,T] \times \mathbb{R} : \tilde{a}_*^{-1}(\omega,t,z) \leq a\} \\ &\quad = \{(\omega,t,z) \in \Omega \times [0,T] \times \mathbb{R} : z \leq \tilde{a}_*(\omega,t,a)\} \\ &\quad = \{(\omega,t,z) \in \Omega \times [0,T] \times \mathbb{R} : z \leq a_*(\omega,t,a) + \gamma a\} \in \mathcal{F} \otimes \mathcal{B}([0,T]) \otimes \mathcal{B}(\mathbb{R}), \end{aligned}$$

because a_* is measurable (see Lemma 3.3). Consequently, the function $\tilde{a}_*^{-1}$ is measurable. Moreover, for all $z \in \mathbb{R}, t \in [0,T], a \in \mathbb{R}$ we have

$$\begin{aligned} &\{(\omega,s) \in \Omega \times [0,t] : \tilde{a}_*^{-1}(\omega,s,z) \leq a\} \\ &\quad = \{(\omega,s) \in \Omega \times [0,t] : z \leq \tilde{a}_*(\omega,t,a) = a_*(\omega,s,a) + \gamma a\} \\ &\quad = \{(\omega,s) \in \Omega \times [0,t] : z - \gamma a \leq a_*(\omega,s,a)\} \in \mathcal{F}_t \otimes \mathcal{B}([0,t]), \end{aligned}$$

since $a_*(\cdot,\cdot,a)$ is progressively measurable for all $a \in \mathbb{R}$ again due to Lemma 3.3.

In order to show that the function $\tilde{a}_*^{-1}$ is Lipschitz continuous in z, let $z_1, z_2 \in \mathbb{R}$ with $z_1 < z_2$ and $(\omega,t) \in \Omega \times [0,T]$ be fixed. We distinguish the following cases:

(i) If $z_1, z_2 \leq l + \gamma f_a(\omega,t,l)$ or $z_1, z_2 \geq r + \gamma f_a(\omega,t,r)$ we have by the definition of $\tilde{a}_*^{-1}$

$$|\tilde{a}_*^{-1}(t,z_1) - \tilde{a}_*^{-1}(t,z_2)| = \frac{1}{\gamma}|z_1 - z_2|.$$

Moreover, we see that

$$\tilde{a}_*^{-1}(t,z_2) - \tilde{a}_*^{-1}(t,z_1) \geq \frac{1}{\gamma + \frac{1}{\delta_l}}(z_2 - z_1).$$

(ii) If $z_1, z_2 \in [l + \gamma f_a(\omega,t,l), r + \gamma f_a(\omega,t,r)]$ we have that

$$\begin{aligned}
|\tilde{a}_*^{-1}(t,z_1) - \tilde{a}_*^{-1}(t,z_2)| &= \left| \left(f_a^{-1}(t,\cdot) + \gamma \mathrm{Id}\right)^{-1}(z_1) - \left(f_a^{-1}(t,\cdot) + \gamma \mathrm{Id}\right)^{-1}(z_2) \right| \\
&= \left| \gamma + \frac{1}{f_{aa}(t, f_a^{-1}(t,\tilde{z}))} \right|^{-1} |z_1 - z_2| \\
&\leq \frac{1}{\gamma + \frac{1}{\delta_u}} |z_1 - z_2| \leq \frac{1}{\gamma} |z_1 - z_2|,
\end{aligned}$$

where $\tilde{z} \in (z_1, z_2)$ exists due to the mean value theorem. Note that for the last inequality we have used that $f_{aa} \leq \delta_u$. In addition, we have by $f_{aa} \geq \delta_l$

$$\tilde{a}_*^{-1}(t,z_2) - \tilde{a}_*^{-1}(t,z_1) \geq \frac{1}{\gamma + \frac{1}{\delta_l}} (z_2 - z_1).$$

(iii) If $z_1 < l + \gamma f_a(\omega,t,l), z_2 > r + \gamma f_a(\omega,t,r)$ we define $\tilde{z}_1 := l + \gamma f_a(\omega,t,l)$ and $\tilde{z}_2 := r + \gamma f_a(\omega,t,r)$, and observe that

$$\begin{aligned}
&|\tilde{a}_*^{-1}(t,z_1) - \tilde{a}_*^{-1}(t,z_2)| \\
&\quad = \tilde{a}_*^{-1}(t,z_2) - \tilde{a}_*^{-1}(t,z_1) \\
&\quad = \tilde{a}_*^{-1}(t,z_2) - \tilde{a}_*^{-1}(t,\tilde{z}_2) + \tilde{a}_*^{-1}(t,\tilde{z}_2) - \tilde{a}_*^{-1}(t,\tilde{z}_1) + \tilde{a}_*^{-1}(t,\tilde{z}_1) - \tilde{a}_*^{-1}(t,z_1) \\
&\quad \leq \frac{1}{\gamma}(z_2 - \tilde{z}_2) + \frac{1}{\gamma}(\tilde{z}_2 - \tilde{z}_1) + \frac{1}{\gamma}(\tilde{z}_1 - z_1) = \frac{1}{\gamma}(z_2 - z_1) = \frac{1}{\gamma}|z_1 - z_2|,
\end{aligned}$$

where we have used (i) and (ii). If we make the term in the last step smaller instead of larger, we obtain by (i) and (ii) again the lower estimate

$$\tilde{a}_*^{-1}(t,z_2) - \tilde{a}_*^{-1}(t,z_1) \geq \frac{1}{\gamma + \frac{1}{\delta_l}} (z_2 - z_1).$$

(iv) In the case of $z_1 \in [l + \gamma f_a(\omega,t,l), r + \gamma f_a(\omega,t,r)]$ and $z_2 > r + \gamma f_a(\omega,t,r)$, or $z_1 < l + \gamma f_a(\omega,t,l)$ and $z_2 \in [l + \gamma f_a(\omega,t,l), r + \gamma f_a(\omega,t,r)]$ one can show the properties analogous to (iii).

We conclude that $\tilde{a}_*^{-1}$ is Lipschitz continuous in z with Lipschitz constant equal to $\frac{1}{\gamma}$. Moreover, we see that (3.6) holds true. $\square$

We observe that the parameters of the FBSDE (3.4) are given by

$$\begin{aligned}
\mu(\omega,t,y) &:= b_t(\omega) + B_t(\omega) y, \\
\sigma(\omega,t,z) &:= z - 2\gamma \tilde{a}_*^{-1}(\omega,t,z), \\
f(\omega,t,x) &:= b_t(\omega) + B_t(\omega) x,
\end{aligned}$$

for $\omega \in \Omega, t \in [0,T]$ and $x, y, z \in \mathbb{R}$. The essential points to consider the auxiliary FBSDE (3.4) instead of our original FBSDE (3.3) are the following two observations.

Proposition 3.6. *The parameters (ξ, μ, σ, f) of the FBSDE (3.4) satisfy SLC and for the Lipschitz constant of σ in z we have $L_{\sigma,z} = 1$.*

Proposition 3.7. *If $(X^x, P^x, \tilde{Z}^x)$ is a solution to the FBSDE (3.4) starting in $x \in \mathbb{R}$, then (M^m, Y^m, Z^m), defined by*

$$\begin{aligned} M_s^m &:= \frac{1}{2}\left(X_s^x + P_s^x\right), \\ Y_s^m &:= \frac{1}{2\gamma}\left(X_s^x - P_s^x\right), \\ Z_s^m &:= -\tilde{a}_*^{-1}\left(s, \tilde{Z}_s^x\right), s \in [0,T], \end{aligned}$$

solves the FBSDE (3.3) with initial value $m \in \mathbb{R}$ such that $M_0^m = m$ a.s.

Proof of Proposition 3.6. Recall the definition of SLC (Definition 2.12). First of all, we need to verify that the parameters are Lipschitz continuous. The functions $\mu(\omega, t, \cdot)$ and $f(\omega, t, \cdot)$ are Lipschitz continuous for all $(\omega, t) \in \Omega \times [0, T]$ due to the linearity in these variables. In order to show that the function σ is also Lipschitz continuous in z, recall Lemma 3.5. We have that $\tilde{a}_*^{-1}$ is Lipschitz continuous with Lipschitz constant $\frac{1}{\gamma}$. Consequently, also σ is Lipschitz continuous in z. To determine a bound for the Lipschitz constant $L_{\sigma,z}$, let $(\omega, t) \in \Omega \times [0, T]$ and $z_1, z_2 \in \mathbb{R}, z_1 < z_2$. Since $\sigma(t, z_2) - \sigma(t, z_1)$ is either positive or negative, we distinguish these two cases:

1. Using that $\tilde{a}_*^{-1}$ has the Lipschitz constant $\frac{1}{\gamma}$ we see that

$$\begin{aligned} |\sigma(t,z_1) - \sigma(t,z_2)| &= \sigma(t,z_1) - \sigma(t,z_2) = z_1 - z_2 + 2\gamma\left(\tilde{a}_*^{-1}(t,z_2) - \tilde{a}_*^{-1}(t,z_1)\right) \\ &\le \left(\frac{2\gamma}{\gamma} - 1\right)(z_2 - z_1) = |z_1 - z_2|. \end{aligned}$$

2. Using the lower bound for $\tilde{a}_*^{-1}$ in (3.6) we observe that

$$\begin{aligned} |\sigma(t,z_1) - \sigma(t,z_2)| &= \sigma(t,z_2) - \sigma(t,z_1) = z_2 - z_1 - 2\gamma\left(\tilde{a}_*^{-1}(t,z_2) - \tilde{a}_*^{-1}(t,z_1)\right) \\ &\le \left(1 - \frac{2\gamma}{\gamma + \frac{1}{\delta_l}}\right)|z_1 - z_2| \le |z_1 - z_2|. \end{aligned}$$

   The fact that $\gamma = \frac{1}{2\delta_u} \le \frac{1}{\delta_u} \le \frac{1}{\delta_l}$ then implies

$$0 < \frac{2\gamma}{\gamma + \frac{1}{\delta_l}} \le \frac{2\gamma}{\frac{1}{\delta_l}} = \frac{\delta_l}{\delta_u} \le 1 \qquad \text{and} \qquad 0 \le 1 - \frac{2\gamma}{\gamma + \frac{1}{\delta_l}} \le 1.$$

Note that in the first case we actually have equality if, for example, $z_1, z_2 < l + \gamma f_a(\omega, t, l)$. Thus $L_{\sigma,z} = 1$ and $L_{\sigma,z}^{-1} = 1$.

Now we turn to Lipschitz continuity of the terminal condition ξ. Let $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2$ and $\omega \in \Omega$. Then according to the mean value theorem there is an $\tilde{x} \in (x_1, x_2)$ such that

$$\begin{aligned} |\xi(x_1) - \xi(x_2)| = |\xi'(\tilde{x})|\,|x_1 - x_2| &= \frac{1 - \gamma g''((\mathrm{Id} + \gamma g')^{-1}(\tilde{x}))}{1 + \gamma g''((\mathrm{Id} + \gamma g')^{-1}(\tilde{x}))}|x_1 - x_2| \\ &\le \frac{1}{1 + \gamma\delta_l}|x_1 - x_2|, \end{aligned}$$

because $g''(\omega, x) \geq \delta_l$ for all $(\omega, x) \in \Omega \times \mathbb{R}$. Consequently, we have for the Lipschitz constant $L_{\xi,x}$ of ξ in x that

$$L_{\xi,x} \leq \frac{1}{1+\gamma\delta_l} < 1 = L_{\sigma,z}^{-1},$$

as required. Finally, we show that the parameters are essentially bounded:

- $\|\mu(\cdot,\cdot,0)\|_\infty + \|f(\cdot,\cdot,0)\|_\infty \leq 2\|b\|_\infty < \infty$, because the process b is bounded.
- If $0 \in [l + \gamma f_a(\omega, t, l), r + \gamma f_a(\omega, t, r)]$ for some (ω, t), we observe that it holds

$$\begin{aligned}
&\left|\tilde{a}_*^{-1}(\omega, t, 0)\right| = \left|(f_a^{-1}(\omega, t, \cdot) + \gamma \mathrm{Id})^{-1}(0)\right| \\
&= \left|(f_a^{-1}(\omega, t, \cdot) + \gamma \mathrm{Id})^{-1}(0) - (f_a^{-1}(\omega, t, \cdot) + \gamma \mathrm{Id})^{-1}\left((f_a^{-1}(\omega, t, \cdot) + \gamma \mathrm{Id})(0)\right)\right| \\
&\leq \frac{1}{\gamma}\left|(f_a^{-1}(\omega, t, \cdot) + \gamma \mathrm{Id})(0)\right| = \frac{1}{\gamma}\left|f_a^{-1}(\omega, t, 0)\right| \\
&= \frac{1}{\gamma}\left|f_a^{-1}(\omega, t, 0) - f_a^{-1}(\omega, t, f_a(\omega, t, 0))\right| \leq \frac{1}{\gamma\delta_l}\left|f_a(\omega, t, 0)\right|,
\end{aligned}$$

  where we have used the Lipschitz continuity of $\tilde{a}_*^{-1}$ (see Lemma 3.5). Consequently, we have by the definition of σ and (C2)

$$\|\sigma(\cdot,\cdot,0)\|_\infty = 2\gamma \operatorname*{ess\,sup}_{(\omega,t)} \left|\tilde{a}_*^{-1}(\omega, t, 0)\right| \leq 2\left(|l| + |r| + \frac{1}{\delta_l}\|f_a(\cdot,\cdot,0)\|_\infty\right) < \infty.$$

- We have by the definition of ξ in (3.5) that

$$\begin{aligned}
|\xi(0)| &= 2\ |\left(\mathrm{Id} + \gamma g'\right)^{-1}(0)| \\
&= 2\left|\left(\mathrm{Id} + \gamma g'\right)^{-1}(0) - \left(\left(\mathrm{Id} + \gamma g'\right)^{-1} \circ \left(\mathrm{Id} + \gamma g'\right)\right)(0)\right| \\
&\leq 2\left(\sup_{x\in\mathbb{R}} \frac{1}{1+\gamma g''(x)}\right)\left|\left(\mathrm{Id} + \gamma g'\right)(0)\right| \leq 2\gamma\|g'(\cdot,0)\|_\infty < \infty, \text{ a.s.},
\end{aligned}$$

  by (C2) and hence $\|\xi(\cdot,0)\|_\infty < \infty$.

To sum up, we have that the parameters (ξ, μ, σ, f) fulfil SLC. □

Proof of Proposition 3.7. Let $(X^x, P^x, \tilde{Z}^x)$ be a solution to the FBSDE (3.4) with initial condition $x \in \mathbb{R}$. This implies, in particular, that $X^x, P^x \in \mathcal{S}^2_{0,T}$ and $\tilde{Z}^x \in \mathcal{H}^2_{0,T}$. Consequently, the random variables X_0^x and P_0^x are $\mathcal{F}_0$-measurable and thus almost surely constant. Thus, M_0^m is also almost surely constant and hence there exists a deterministic initial value $m \in \mathbb{R}$ such that $M_0^m = m$ a.s.

The processes M^m, Y^m belong to the space $\mathcal{S}^2_{0,T}$ since they are only linear combinations of X^x and P^x. The process Z^m is progressively measurable because it is a composition of progressively measurable functions. In other words, $\tilde{a}_*^{-1}$ is measurable, $\tilde{a}_*^{-1}(\cdot,\cdot,z)$ is progressively measurable and $\tilde{a}_*^{-1}$ is continuous in z due to Lemma 3.5. Thus, the composition with the progressively measurable process $\tilde{Z}^x$ is progressively measurable, which can be shown as in the proof of Lemma 3.3. In addition, we have $Z^m \in \mathcal{H}^2_{0,T}$ because

$$\begin{aligned}
\mathbb{E}\left[\int_0^T (Z^m_s)^2\,\mathrm{d}s\right] &= \mathbb{E}\left[\int_0^T (\tilde{a}_*^{-1}(s,\tilde{Z}^x_s))^2\,\mathrm{d}s\right] \\
&= \mathbb{E}\left[\int_0^T \left(\tilde{a}_*^{-1}(s,\tilde{Z}^x_s) - \tilde{a}_*^{-1}(s,0) + \tilde{a}_*^{-1}(s,0)\right)^2\,\mathrm{d}s\right] \\
&\le \mathbb{E}\left[\int_0^T \left(\frac{2}{\gamma^2}\left(\tilde{Z}^x_s\right)^2 + 2\tilde{a}_*^{-1}(s,0)^2\right)\,\mathrm{d}s\right] \\
&\le \mathbb{E}\left[\int_0^T \left(\frac{2}{\gamma^2}\left(\tilde{Z}^x_s\right)^2 + 2\|\tilde{a}_*^{-1}(\cdot,\cdot,0)\|^2_\infty\right)\,\mathrm{d}s\right] < \infty.
\end{aligned}$$

Now it is straightforward to verify that the processes M^m, Y^m, Z^m satisfy the FBSDE (3.3). Note that all the integrals in (3.3) are well-defined. Writing X^x and P^x as forward processes, i.e.

$$\begin{aligned}
dX^x_s &= (b_s + B_s P^x_s)\,ds + \left(\tilde{Z}^x_s - 2\gamma\tilde{a}_*^{-1}(s,\tilde{Z}^x_s)\right) dW_s, \\
dP^x_s &= (b_s + B_s X^x_s)\,ds + \tilde{Z}^x_s dW_s,
\end{aligned}$$

we observe that (M^m, Y^m, Z^m) fulfils (3.3), because for all $s \in [0,T]$ we have:

$$\begin{aligned}
dM^m_s &= \frac{1}{2}\left(dX^x_s + dP^x_s\right) \\
&= \frac{1}{2}\left[\,(b_s + B_s P^x_s) + (b_s + B_s X^x_s)\,\right]ds + \frac{1}{2}\left[\left(\tilde{Z}^x_s - 2\gamma\tilde{a}_*^{-1}\left(s,\tilde{Z}^x_s\right)\right) + \tilde{Z}^x_s\right] dW_s \\
&= (b_s + B_s M^m_s)\,ds + \left(\tilde{Z}^x_s - \gamma\tilde{a}_*^{-1}\left(s,\tilde{Z}^x_s\right)\right) dW_s \\
&= (b_s + B_s M^m_t)\,ds + (\tilde{a}_*(s,-Z^m_s) + \gamma Z^m_s)\,dW_s \\
&= (b_s + B_s M^m_t)\,ds + a_*(s,-Z^m_s)dW_s,
\end{aligned}$$

and

$$\begin{aligned}
dY^m_s &= \frac{1}{2\gamma}\left(dX^x_s - dP^x_s\right) \\
&= \frac{1}{2\gamma}\left[b_s + B_s P^x_s - (b_s + B_s X^x_s)\,\right]ds + \frac{1}{2\gamma}\left[\tilde{Z}^x_s - 2\gamma\tilde{a}_*^{-1}\left(s,\tilde{Z}^x_s\right) - \tilde{Z}^x_s\right] dW_s \\
&= -B_s Y^m_s ds + Z^m_s dW_s.
\end{aligned}$$

We calculate the terminal condition of (3.3):

$$\begin{aligned}
Y^m_T &= \frac{1}{2\gamma}(X^x_T - P^x_T) = \frac{1}{2\gamma}(X^x_T - \xi(X^x_T)) = \frac{1}{2\gamma}\left(X^x_T - 2(\mathrm{Id}+\gamma g')^{-1}(X^x_T) + X^x_T\right) \\
&= \frac{1}{\gamma}\left(X^x_T - (\mathrm{Id}+\gamma g')^{-1}(X^x_T)\right) \\
&= \frac{1}{\gamma}\left[(\mathrm{Id}+\gamma g')\left((\mathrm{Id}+\gamma g')^{-1}(X^x_T)\right) - (\mathrm{Id}+\gamma g')^{-1}(X^x_T)\right] = g'\left((\mathrm{Id}+\gamma g')^{-1}(X^x_T)\right), \text{a.s.}
\end{aligned}$$

Moreover, we have

$$\begin{aligned}(\mathrm{Id}+\gamma g')^{-1}(X_T^x) &= \frac{1}{2}\left(X_T^x + \left(2(\mathrm{Id}+\gamma g')^{-1}(X_T^x) - X_T^x\right)\right)\\ &= \frac{1}{2}\left(X_T^x + \xi(X_T^x)\right) = \frac{1}{2}\left(X_T^x + P_T^x\right) = M_T^m, \text{ a.s.,}\end{aligned}$$

which implies $Y_T^m = g'(M_T^m)$ a.s. Finally, the triple (M^m, Y^m, Z^m) solves the FBSDE (3.3) with initial condition m and terminal condition $g'(M_T^m)$. □

3.4 Applying the method of decoupling fields

We have proven that we can recover a solution to our original FBSDE (3.3) from the auxiliary FBSDE (3.4). In this section we employ the method of decoupling fields to solve the latter equation. Note that the parameters of the FBSDE (3.4) fulfil SLC as we have shown in Proposition 3.6. Consequently, according to Theorem 2.21 there exists a strongly regular decoupling field

$$u : \Omega \times I_{max} \times \mathbb{R} \to \mathbb{R}$$

on the maximal interval $I_{max} \subseteq [0,T]$ such that $L_{u,x} < L_{\sigma,z}^{-1}$. We can assume that $u(\omega,t,\cdot)$ is Lipschitz continuous and weakly differentiable for all $(\omega,t) \in \Omega \times I_{max}$ (see Remark 2.14). We denote by u_x a version of the weak derivative of u that is equal to the classical derivative of u, where it exists, and equal to zero otherwise. The Lipschitz continuity of u especially implies that u is almost everywhere classically differentiable and thus $u_x(\omega,t,x)$ is identical to the classical derivative for almost all $x \in \mathbb{R}$ and all $(\omega,t) \in \Omega \times I_{max}$.

Our goal in this section is to show that only $I_{max} = [0,T]$ is possible. For this purpose we prove that

$$\lim_{t \downarrow t_{min}} L_{u(t,\cdot),x} \neq L_{\sigma,z}^{-1}, \tag{3.7}$$

for all $t_{min} \in [0,T]$. Then Proposition 2.22 implies that $I_{max} = (t_{min}, T]$ cannot hold true and thus $I_{max} = [0,T]$ by Theorem 2.21.

From now on, let $t_0 \in I_{max}$ be fixed. As a result of the strong regularity of u, there exists a solution $(X^x, P^x, \tilde{Z}^x)$ to the FBSDE (3.4) on $[t_0,T]$ for the deterministic initial value $x \in \mathbb{R}$. It holds, in particular, that

$$P_t^x = u(t, X_t^x), \text{ a.s.,}$$

for all $t \in [t_0,T]$ and $x \in \mathbb{R}$. In addition, X_t^x and P_t^x are weakly differentiable w.r.t. x for all $t \in [t_0,T]$ and the process $\tilde{Z}^x$ is weakly differentiable w.r.t. x. We define $(\partial_x X^x, \partial_x P^x, \partial_x \tilde{Z}^x)$ as a version of the weak derivative of $(X^x, P^x, \tilde{Z}^x)$ such that $(\partial_x X_s^x, \partial_x P_s^x)$ is a weak derivative of (X_s^x, P_s^x) for all $s \in [t_0,T]$. In order to obtain the dynamics of $(\partial_x X^x, \partial_x P^x, \partial_x \tilde{Z}^x)$, we differentiate the forward and backward equation of the FBSDE (3.4).

Lemma 3.8. *There exists a process $\Sigma : \Omega \times [t_0, T] \times \mathbb{R} \to \mathbb{R}$ that is bounded by $L_{\sigma,z}$ and fulfils*

$$\partial_x \left(\sigma(s, \tilde{Z}_s) \right) = \Sigma_s \partial_x \tilde{Z}_s$$

for $\mathbb{P} \otimes \lambda \otimes \lambda$-almost every $(\omega, s, x) \in \Omega \times [t_0, T] \times \mathbb{R}$. Moreover, the weak derivative $(\partial_x X^x, \partial_x P^x, \partial_x \tilde{Z}^x)$ satisfies for all $t \in [t_0, T]$ the FBSDE

$$\begin{aligned} \partial_x X_t^x &= 1 + \int_{t_0}^t B_s \partial_x P_s^x \, \mathrm{d}s + \int_{t_0}^t \Sigma_s \partial_x \tilde{Z}_s^x \, \mathrm{d}W_s, \\ \partial_x P_t^x &= \xi'(X_T^x) \partial_x X_T^x - \int_t^T B_s \partial_x X_s^x \, \mathrm{d}s - \int_t^T \partial_x \tilde{Z}_s^x \, \mathrm{d}s, \end{aligned} \tag{3.8}$$

for $\mathbb{P} \otimes \lambda$-almost all $(\omega, x) \in \Omega \times \mathbb{R}$.

Proof. The right hand sides of the FBSDE (3.4) are weakly differentiable w.r.t. x since X_t^x and P_t^x are weakly differentiable for all $t \in [t_0, T]$. However, we calculate versions of those weak derivatives that are equal to the right hand sides of (3.8). To that end, we note that:

(i) For all $s \in [t_0, T]$ the functions $\mu(s, P_s^x)$ and $f(s, X_s^x)$ are weakly differentiable a.s. with weak derivatives

$$\begin{aligned} \partial_x \mu(s, P_s^x) &= B_s \partial_x P_s^x, \text{ a.s.,} \\ \partial_x f(s, X_s^x) &= B_s \partial_x X_s^x, \text{ a.s.,} \end{aligned}$$

for almost all $x \in \mathbb{R}$. Let now $t \in [t_0, t]$ be fixed. We have that

$$\begin{aligned} &\mathbb{E}\left[\int_{t_0}^t (|\mu(s, P_s^x)| + |f(s, X_s^x)|) \, \mathrm{d}s \right] \\ &\qquad \leq \mathbb{E}\left[\int_{t_0}^t (L(|X_s^x| + |P_s^x|) + \|\mu(\cdot, \cdot, 0)\|_\infty + \|f(\cdot, \cdot, 0)\|_\infty) \, \mathrm{d}s \right] < \infty, \end{aligned}$$

for all $x \in \mathbb{R}$, because $X^x, P^x \in \mathcal{S}^2_{t_0,T}$ and the parameters satisfy SLC with Lipschitz constant $L \geq 0$. Additionally, we observe that

$$\begin{aligned} &\operatorname*{ess\,sup}_{x \in \mathbb{R}} \mathbb{E}\left[\int_{t_0}^t (|\partial_x \mu(s, P_s^x)| + |\partial_x f(s, X_s^x)|) \, \mathrm{d}s \right] \\ &\qquad \leq \|B\|_\infty \operatorname*{ess\,sup}_{x \in \mathbb{R}} \mathbb{E}\left[\int_{t_0}^t (|\partial_x X_s^x| + |\partial_x P_s^x|) \, \mathrm{d}s \right] < \infty, \end{aligned}$$

due to the strong regularity of the decoupling field u. The conditions (1)-(3) of Proposition 2.36 are therefore fulfilled and hence the mappings

$$x \mapsto \int_{t_0}^t \mu(s, P_s^x) \, \mathrm{d}s \qquad \text{and} \qquad x \mapsto \int_{t_0}^t f(s, X_s^x) \, \mathrm{d}s$$

are weakly differentiable for all fixed $t \in [t_0, T]$ with weak derivatives

$$\begin{aligned} \partial_x \int_{t_0}^t \mu(s, P_s^x) \, \mathrm{d}s &= \int_{t_0}^t B_s \partial_x P_s^x \, \mathrm{d}s, \text{ a.s., and} \\ \partial_x \int_{t_0}^t f(s, X_s^x) \, \mathrm{d}s &= \int_{t_0}^t B_s \partial_x X_s^x \, \mathrm{d}s, \text{ a.s.,} \end{aligned}$$

respectively, for almost all $x \in \mathbb{R}$.

(ii) σ is Lipschitz continuous in z with Lipschitz constant $L_{\sigma,z}$, but not everywhere classically differentiable. Thus, we apply Proposition 2.41 to $\sigma(s, \tilde{Z}_s^x)$, which yields the weak differentiability and the existence of a process $\Sigma : \Omega \times [t_0, T] \times \mathbb{R} \to \mathbb{R}$ that is bounded by $L_{\sigma,z}$ and satisfies

$$\partial_x \left(\sigma(s, \tilde{Z}_s^x) \right) = \Sigma(s, x) \partial_x \tilde{Z}_s^x$$

for $\mathbb{P} \otimes \lambda \otimes \lambda$-almost every $(\omega, s, x) \in \Omega \times [t_0, T] \times \mathbb{R}$. In the following we abbreviate $\Sigma(s, x)$ by Σ_s. Now we observe for all $t \in [t_0, T]$ and $x \in \mathbb{R}$ that

$$\mathbb{E}\left[\int_{t_0}^{t} |\sigma(s, Z_s^x)|^2 \,\mathrm{d}s \right] \leq \mathbb{E}\left[\int_{t_0}^{t} \left(2L_{\sigma,z}^2 |Z_s^x|^2 + 2\|\sigma(\cdot, \cdot, 0)\|_\infty^2 \right) \,\mathrm{d}s \right] < \infty,$$

because $\tilde{Z}^x \in \mathcal{H}_{t_0,T}^2$ and the parameters of the FBSDE (3.4) satisfy SLC. Moreover, it holds true that for all $t \in [t_0, T]$

$$\operatorname*{ess\,sup}_{x \in \mathbb{R}} \mathbb{E}\left[\int_{t_0}^{t} |\Sigma_s \partial_x \tilde{Z}_s^x|^2 \,\mathrm{d}s \right] \leq L_{\sigma,z}^2 \operatorname*{ess\,sup}_{x \in \mathbb{R}} \mathbb{E}\left[\int_{t_0}^{t} |\partial_x \tilde{Z}_s^x|^2 \,\mathrm{d}s \right] < \infty,$$

since the decoupling field u is strongly regular. Hence the conditions of Proposition 2.37 are fulfilled and thus the mapping $x \mapsto \int_{t_0}^{t} \sigma(s, \tilde{Z}_s^x) \,\mathrm{d}W_s$ is weakly differentiable for all fixed $t \in [t_0, T]$ with weak derivative

$$\partial_x \int_{t_0}^{t} \sigma(s, \tilde{Z}_s^x) \,\mathrm{d}W_s = \int_{t_0}^{t} \Sigma_s \partial_x \tilde{Z}_s^x \,\mathrm{d}W_s,$$

for $\mathbb{P} \otimes \lambda$-almost all $(\omega, x) \in \Omega \times \mathbb{R}$.

(iii) We know that ξ is Lipschitz continuous and everywhere continuously differentiable for all $\omega \in \Omega$. Therefore, we obtain by Proposition 2.38 that $\xi(X_T^x)$ is weakly differentiable w.r.t. x and

$$\partial_x \left(\xi(X_T^x) \right) = \xi'(X_T^x) \partial_x X_T^x$$

for $\mathbb{P} \otimes \lambda$-almost all $(\omega, x) \in \Omega \times \mathbb{R}$.

The properties (i)-(iii) imply that for all $t \in [t_0, T]$ the right hand sides of (3.4) have versions of the weak derivatives that are equal to

$$\begin{gathered} 1 + \int_{t_0}^{t} B_s \partial_x P_s^x \,\mathrm{d}s + \int_{t_0}^{t} \Sigma_s \partial_x \tilde{Z}_s^x \,\mathrm{d}W_s, \text{ and} \\ \xi'(X_T^x) \partial_x X_T^x - \int_{t}^{T} B_s \partial_x X_s^x \,\mathrm{d}s - \int_{t}^{T} \partial_x \tilde{Z}_s^x \,\mathrm{d}s, \end{gathered}$$

respectively, due to Proposition 2.31. But these expressions are also versions of the weak derivatives of X_t^x and P_t^x and thus for all $t \in [t_0, T]$ equation (3.8) holds true for $\mathbb{P} \otimes \lambda$-almost all $(\omega, x) \in \Omega \times \mathbb{R}$. In other words, the properties (i)-(iii) allow us to interchange weak differentiation and the integrals in the FBSDE (3.4), which implies (3.8). $\square$

Remark 3.9. Lemma 3.8 implies that there are versions of the weak derivatives $\partial_x X^x$ and $\partial_x P^x$ that are continuous in time for all $(\omega, x) \in \Omega \times \mathbb{R}$. One can just consider the right hand sides of (3.8) as processes and note that they are continuous versions of the weak derivatives.

From now on, we assume that $\partial_x X^x$ and $\partial_x P^x$ are continuous in time. In order to calculate a bound for the Lipschitz constant of u, we study the so-called *gradient process* V, defined by

$$V_t^x(\omega) := u_x(\omega, t, X_t^x(\omega)),$$

for $(\omega, t, x) \in \Omega \times [t_0, T] \times \mathbb{R}$. The chain rule for weak derivatives in Proposition 2.38 implies that for any fixed $t \in [t_0, T]$

$$\partial_x P_t^x = u_x(t, X_t^x)\partial_x X_t^x = V_t^x \partial_x X_t^x, \tag{3.9}$$

for $\mathbb{P} \otimes \lambda$-almost all $(\omega, x) \in \Omega \times \mathbb{R}$.

For the following considerations we choose a fixed $x \in \mathbb{R}$ such that:

(1) $\partial_x X_{t_0}^x = 1$ a.s.,

(2) $\partial_x P_{t_0}^x = V_{t_0}^x \partial_x X_{t_0}^x$ a.s.,

(3) (3.8) holds true for $\mathbb{P} \otimes \lambda$-almost all $(\omega, t) \in \Omega \times [t_0, T]$,

(4) (3.9) holds true for $\mathbb{P} \otimes \lambda$-almost all $(\omega, t) \in \Omega \times [t_0, T]$,

(5) $\partial_x X^x, \partial_x P^x \in \mathcal{S}_{t_0,T}^2, \partial_x \tilde{Z}^x \in \mathcal{H}_{t_0,T}^2$.

Note that we have chosen x outside of a Lebesgue null set. In particular, (5) holds true because of the strong regularity of u (see (2.7)). Moreover, we emphasize that under these assumptions equation (3.8) is even true for all $t \in [t_0, T]$ and almost all $\omega \in \Omega$ since $\partial_x X^x$ and $\partial_x P^x$ are continuous in time. In the following we omit the initial value x in the superscript of the processes that we consider.

Lemma 3.10. *The gradient process V is bounded, i.e. there exists a constant $K > 0$ such that $|V_t| \leq K$ for all $(\omega, t) \in \Omega \times [t_0, T]$. We can choose this constant such that*

$$K = \sup_{s \in [t_0, T]} L_{u(s,\cdot),x} < L_{\sigma,z}^{-1} = 1. \tag{3.10}$$

We point out that the constant K might depend on the fixed time t_0. However, we prove later on that $L_{u(t,\cdot),x}$ can be uniformly bounded away from 1.

Proof. We know that the function $u(\omega, t, \cdot)$ is almost everywhere differentiable for all $(\omega, t) \in \Omega \times [t_0, T]$, because of the Lipschitz continuity of u. We have chosen the version u_x of the weak derivative of u such that it is equal to the classical derivative where it exists, and equal to zero otherwise. Therefore, we have for all $(\omega, t) \in \Omega \times [t_0, T]$ and almost all $x' \in \mathbb{R}$ that

$$|u_x(t, x')| = \lim_{h \to 0} \frac{|u(t, x' + h) - u(t, x')|}{|h|} \leq \sup_{s \in [t_0, T]} L_{u(s,\cdot),x} < L_{\sigma,z}^{-1},$$

because of the Lipschitz continuity and regularity of u. Consequently, we observe for all $(\omega, t) \in \Omega \times [t_0, T]$ and $x' \in \mathbb{R}$ that

$$|u_x(t, x')| \leq \sup_{s \in [t_0, T]} L_{u(s,\cdot),x} < L_{\sigma,z}^{-1}.$$

Then we have for the gradient process V

$$|V_t| = |u_x(t, X_t)| \leq \sup_{s \in [t_0, T]} L_{u(s,\cdot),x} =: K,$$

for all $(\omega, t) \in \Omega \times [t_0, T]$. □

We now introduce *BMO-processes*, which we will use to construct an equivalent probability measure in Lemma 3.13. Note that details and important statements about these processes can be found, for instance, in [6] and [7].

Definition 3.11. A progressively measurable process Z is called *BMO-process* if there exists a $C \geq 0$ such that for all $t \in [0, T]$

$$\mathbb{E}\left[\int_t^T Z_s^2 \, \mathrm{d}s \middle| \mathcal{F}_t\right] \leq C, \text{ a.s.}$$

Next we show that the gradient process V has a version that fulfils a BSDE. This will later on help us to improve the bound of V.

Lemma 3.12. *The process $V = (V_t)_{t \in [t_0, T]}$ has a time-continuous version $\hat{V} = (\hat{V}_t)_{t \in [t_0, T]}$ that is an Itô process. Moreover, there exists a process $\hat{Z} \in \mathcal{H}^2_{t_0,T}$ such that $(\hat{V}, \hat{Z})$ solves the BSDE*

$$\hat{V}_t = \xi'(X_T) - \int_t^T \rho(s, \hat{V}_s, \hat{Z}_s) \, \mathrm{d}s - \int_t^T \hat{Z}_s \, \mathrm{d}W_s, \ t \in [t_0, T], \tag{3.11}$$

where $\rho : \Omega \times [t_0, T] \times [-K, K] \times \mathbb{R} \to \mathbb{R}$ is given by

$$\rho(t, v, z) = B_t(1 - v)(v + 1) - \frac{\Sigma_t z^2}{1 - \Sigma_t v},$$

and K is defined as in (3.10)*. $(\hat{V}, \hat{Z})$ is the unique solution of BSDE* (3.11) *with bounded first component. Furthermore, the process $\hat{Z}$ is a BMO-process.*

Proof. First of all, we show that ρ is well-defined, i.e. the denominator $1 - \Sigma_t v$ is bounded away from zero. Let $v \in [-K, K]$. Since Σ is bounded by $L_{\sigma,z}$ and $|v| \leq K$, we have $|\Sigma_t v| \leq L_{\sigma,z} K < L_{\sigma,z} L_{\sigma,z}^{-1} = 1$, and consequently,

$$|1 - \Sigma_t v| \geq 1 - L_{\sigma,z} K > 0.$$

Now we prove the existence of the time-continuous version $\hat{V}$. Let $n \in \mathbb{N}$ and $\tau_n := \inf\left\{t \geq t_0 : \partial_x X_t \leq \frac{1}{n}\right\} \wedge T$. Because $\partial_x X_{t_0} = 1$ we know by (3.9) and our choice of x that for $\mathbb{P}$-almost all $\omega \in \Omega$

$$V_t = \frac{\partial_x P_t}{\partial_x X_t} = \partial_x P_t \frac{1}{\partial_x X_t}, \text{ for a.a. } t \in [t_0, \tau_n].$$

Applying Itô's formula to the right hand side of the equation implies that there exists a version of V that is almost surely continuous and an Itô process on $[t_0, \tau_n]$. We denote this version by

$$\hat{V}_t := u_x(t_0, x) + \int_{t_0}^{t} \eta_s \,\mathrm{d}s + \int_{t_0}^{t} \hat{Z}_s \,\mathrm{d}W_s, \ t \in [t_0, \tau_n], \tag{3.12}$$

for some suitable processes η and $\hat{Z}$. In particular, we mean by version here that for almost all $\omega \in \Omega$ we have $V_t = \hat{V}_t$ for almost every $t \in [t_0, \tau_n]$. We point out that although V and $\hat{V}$ are just almost everywhere equal, we have that $V_{t_0} = \hat{V}_{t_0}$ a.s. due to the choice of x and the definition of $\hat{V}$. Moreover, note that the bound K for the process V can be transferred to the version $\hat{V}$, because one can argue as follows for almost all $\omega \in \Omega$. Let $t \in [t_0, \tau_n]$. Then for all $\varepsilon > 0$ there exists a time $t' \in [t_0, \tau_n]$ such that $\hat{V}_{t'} = V_{t'}$ and thus

$$|\hat{V}_t| \leq |\hat{V}_t - \hat{V}_{t'}| + |\hat{V}_{t'}| \leq \varepsilon + K,$$

by the continuity of $\hat{V}$ and the boundedness of V according to Lemma 3.10. The arbitrariness of t and ε implies that for almost all $\omega \in \Omega$ it holds

$$|\hat{V}_t| \leq K, \ t \in [t_0, \tau_n]. \tag{3.13}$$

Now we apply the product formula to $\hat{V}_t \partial_x X_t$ for $t \in [t_0, \tau_n]$ to determine the processes η and $\hat{Z}$. We obtain that

$$d\left(\hat{V}_t \partial_x X_t\right) = \left[\hat{V}_t B_t \partial_x P_t + \partial_x X_t \eta_t + \hat{Z}_t \Sigma_t \partial_x \tilde{Z}_t\right] dt + \left[\hat{Z}_t \partial_x X_t + \hat{V}_t \Sigma_t \partial_x \tilde{Z}_t\right] dW_t. \tag{3.14}$$

This expression has to coincide with the dynamics of $\partial_x P_t$ presented in (3.8), because for almost all $\omega \in \Omega$ we have

$$\partial_x P_t = \hat{V}_t \partial_x X_t, \ t \in [t_0, \tau_n].$$

Consequently, we can determine η and $\hat{Z}$ by comparing the drift and diffusion terms of (3.8) and (3.14). To be more precise, we choose η and $\hat{Z}$ such that

$$\hat{V}_t B_t \partial_x P_t + \partial_x X_t \eta_t + \hat{Z}_t \Sigma_t \partial_x \tilde{Z}_t = B_t \partial_x X_t, \tag{3.15}$$

$$\partial_x \tilde{Z}_t = \hat{Z}_t \partial_x X_t + \hat{V}_t \Sigma_t \partial_x \tilde{Z}_t, \tag{3.16}$$

for $t \in [t_0, \tau_n]$, $\mathbb{P}$-almost surely. This means that for almost all $\omega \in \Omega$ and all $t \in [t_0, \tau_n]$:

- By equation (3.16) we obtain that

$$\hat{Z}_t = \frac{1 - \hat{V}_t \Sigma_t}{\partial_x X_t} \partial_x \tilde{Z}_t \qquad \text{and} \qquad \partial_x \tilde{Z}_t = \frac{\hat{Z}_t \partial_x X_t}{1 - \hat{V}_t \Sigma_t}.$$

  Note that the denominator $1 - \hat{V}_t \Sigma_t$ is bounded away from zero since $|\hat{V}_t| \leq K$.

- By equation (3.15), the property $\hat{V}_t \partial_x X_t = \partial_x P_t$ and the above identity for $\partial_x \tilde{Z}_t$ we have that

$$B_t \hat{V}_t^2 \partial_x X_t + \partial_x X_t \eta_t + \hat{Z}_t \Sigma_t \frac{\hat{Z}_t \partial_x X_t}{1 - \hat{V}_t \Sigma_t} = B_t \partial_x X_t,$$

  and hence

$$\eta_t = B_t (1 - \hat{V}_t)(1 + \hat{V}_t) - \frac{\Sigma_t \hat{Z}_t^2}{1 - \hat{V}_t \Sigma_t} = \rho(t, \hat{V}_t, \hat{Z}_t),$$

  because $\partial_x X_t \neq 0$ on $[t_0, \tau_n]$.

Note that $(\hat{V}, \hat{Z})$ can be viewed as a solution of the quadratic BSDE

$$\hat{V}_t = \hat{V}_{\tau_n} - \int_t^{\tau_n} \rho(s, \hat{V}_s, \hat{Z}_s)\,\mathrm{d}s - \int_t^{\tau_n} \hat{Z}_s\,\mathrm{d}W_s$$

on $[t_0, \tau_n]$. This implies that $\hat{Z}$ is a BMO-process on $[t_0, \tau_n]$ according to Theorem A.1.11 in [6], because:

- $\hat{V}$ is bounded,
- the generator can be estimated as follows:

$$\begin{aligned}
-\rho(s, \hat{V}_s, \hat{Z}_s) &= -B_t(1 - \hat{V}_t)(1 + \hat{V}_t) + \frac{\Sigma_s \hat{Z}_t^2}{1 - \hat{V}_t \Sigma_t} \\
&\leq 2\|B\|_\infty (1 + K) + \frac{L_{\sigma,z}}{1 - K L_{\sigma,z}} \hat{Z}_t^2,
\end{aligned}$$

  since $B, \hat{V}, \Sigma$ are bounded by $\|B\|_\infty, K, L_{\sigma,z}$, respectively,
- $\hat{Z}$ is square-integrable on $[t_0, \tau_n]$, since for almost all $\omega \in \Omega$

$$\int_{t_0}^{\tau_n} \hat{Z}_s^2\,\mathrm{d}s = \int_{t_0}^{\tau_n} \left(\frac{1 - \hat{V}_s \Sigma_s}{\partial_x X_s} \right)^2 \left(\partial_x \tilde{Z}_s \right)^2 \mathrm{d}s \leq n^2 (1 + K L_{\sigma,z})^2 \int_{t_0}^{\tau_n} \left(\partial_x \tilde{Z}_s \right)^2 \mathrm{d}s < \infty,$$

  where we have used the definition of τ_n, $|1 - \hat{V}_s \Sigma_s| \leq 1 + |\hat{V}_s \Sigma_s| \leq 1 + K L_{\sigma,z}$ and the choice of x, which ensures that

$$\int_{t_0}^{T} \left(\partial_x \tilde{Z}_s \right)^2 \mathrm{d}s < \infty, \text{ a.s.}$$

We emphasize that the reasoning above does not depend on the choice of $n \in \mathbb{N}$. As a matter of fact, Theorem A.1.11 in [6] even tells us that $\hat{Z}$ has a BMO-norm independent of n, i.e. there exists a constant $C > 0$ such that for all $n \in \mathbb{N}$ and $t \in [t_0, T]$

$$\mathbb{E}\left[\int_t^T \mathbb{1}_{[t_0, \tau_n]}(s) \hat{Z}_s^2\,\mathrm{d}s \middle| \mathcal{F}_t \right] \leq C, \text{ a.s.} \tag{3.17}$$

The constant C only depends on the constants $\|B\|_\infty, K, L_{\sigma,z}$, which appear in the estimates above. Consequently, passing to the limit as $n \to \infty$ yields that $\hat{Z}$ is a BMO-process on $[t_0, \tau)$, where $\tau := \lim_{n\to\infty} \tau_n$. To be more precise, for all $t \in [t_0, T]$ monotone convergence implies

$$\mathbb{E}\left[\int_t^T \mathbb{1}_{[t_0, \tau)}(s) \hat{Z}_s^2\,\mathrm{d}s \middle| \mathcal{F}_t \right] = \lim_{n\to\infty} \mathbb{E}\left[\int_t^T \mathbb{1}_{[t_0, \tau_n]}(s) \hat{Z}_s^2\,\mathrm{d}s \middle| \mathcal{F}_t \right] \leq C, \text{ a.s.}$$

Now we aim at showing that $\tau = T$ holds true almost surely. To that end, we observe that $\partial_x X$ satisfies on $[t_0, \tau)$ the linear SDE

$$d\partial_x X_t = \alpha_t \partial_x X_t dt + \beta_t \partial_x X_t dW_t,$$

where

$$\alpha_t := B_t \hat{V}_t \qquad \text{and} \qquad \beta_t := \frac{\Sigma_t \hat{Z}_t}{1 - \hat{V}_t \Sigma_t},$$

which can be verified using the dynamics of $\partial_x X_t$ in (3.8) and the property $\partial_x P_t = \hat{V}_t \partial_x X_t$. Thus, $\partial_x X$ has the form

$$\partial_x X_{t\wedge\tau_n} = \exp\left(\int_{t_0}^{t\wedge\tau_n} \left(\alpha_s - \frac{1}{2}\beta_s^2\right) \mathrm{d}s + \int_{t_0}^{t\wedge\tau_n} \beta_s \, \mathrm{d}W_s\right), \ t \in [t_0, T], \ n \in \mathbb{N}.$$

We observe that α is bounded, since $|\alpha_t| \le K\|B\|_\infty$, $t \in [t_0, \tau)$, a.s., and that β is bounded by $\hat{Z}$, because

$$|\beta_t| \le \frac{L_{\sigma,z}}{1 - KL_{\sigma,z}} \hat{Z}_t, \ t \in [t_0, \tau), \ \text{a.s.} \tag{3.18}$$

Furthermore, for all $n \in \mathbb{N}$

$$\begin{aligned} \partial_x X_{\tau_n} &\ge \exp\left(-\int_{t_0}^{\tau_n} K\|B\|_\infty \, \mathrm{d}s\right) \exp\left(\int_{t_0}^{\tau_n} \beta_s \, \mathrm{d}W_s - \int_{t_0}^{\tau_n} \frac{1}{2}\beta_s^2 \, \mathrm{d}s\right) \\ &\ge \exp\left(-K\|B\|_\infty T\right) \mathcal{E}_{\tau_n}(M), \ \text{a.s.}, \end{aligned} \tag{3.19}$$

where $M_\cdot := \int_{t_0}^{\cdot\wedge\tau} \beta_s \, \mathrm{d}W_s$ and $\mathcal{E}(M)$ is the exponential of M, i.e.

$$\mathcal{E}_t(M) = \exp\left(\int_{t_0}^{t\wedge\tau} \beta_s \, \mathrm{d}W_s - \int_{t_0}^{t\wedge\tau} \frac{1}{2}\beta_s^2 \, \mathrm{d}s\right), \ t \in [t_0, T].$$

Note that M is a martingale on $[t_0, T]$, because

$$\mathbb{E}\left[\int_{t_0}^{T} \mathbb{1}_{[t_0,\tau)}(s)\beta_s^2 \, \mathrm{d}s\right] \le \mathbb{E}\left[\int_{t_0}^{\tau} \left(\frac{L_{\sigma,z}}{1 - KL_{\sigma,z}}\right)^2 \hat{Z}_s^2 \, \mathrm{d}s\right] < \infty.$$

The last integral is finite since $\hat{Z}$ is a BMO-process on $[t_0, \tau)$, which implies, in particular, that $\mathbb{E} \int_{t_0}^{\tau} \hat{Z}_s^2 \, \mathrm{d}s < \infty$. Then we have

$$\begin{aligned} \mathbb{P}\left(\lim_{n\to\infty} \partial_x X_{\tau_n} = 0\right) &\le \mathbb{P}\left(\lim_{n\to\infty} \mathcal{E}_{\tau_n}(M) = 0\right) \\ &= \mathbb{P}\left(\lim_{t\to\infty} \mathcal{E}_t(M) = 0\right) \\ &= \mathbb{P}\left(\lim_{t\to\infty} \int_{t_0}^{t\wedge\tau} \frac{1}{2}\beta_s^2 \, \mathrm{d}s = \infty\right) \\ &\le \mathbb{P}\left(\lim_{t\to\infty} \int_{t_0}^{t\wedge\tau} \frac{1}{2}\left(\frac{L_{\sigma,z}}{1 - KL_{\sigma,z}}\right)^2 \hat{Z}_s^2 \, \mathrm{d}s = \infty\right) \\ &= \mathbb{P}\left(\int_{t_0}^{\tau} \hat{Z}_s^2 \, \mathrm{d}s = \infty\right) = 0. \end{aligned}$$

The last probability vanishes because $\hat{Z}$ is a BMO-process on $[t_0, \tau)$, which especially means that the integral in the last probability is finite a.s. For the above calculation we have used the following:

- The first inequality holds true because of (3.19),
- the second equality holds true because $\mathcal{E}$ is continuous and

$$\mathcal{E}_t(M) = \mathcal{E}_\tau(M) = \lim_{n\to\infty} \mathcal{E}_{\tau_n}(M), \text{ a.s., } t \geq T,$$

- the third equality is true due to the statement

$$\mathbb{P}\left(\lim_{t\to\infty} \mathcal{E}_t(M) = 0\right) = \mathbb{P}\left(\lim_{t\to\infty} \langle M, M\rangle_t = \infty\right),$$

  which can be found in [12, p. 157], and
- the last two steps are possible since β is bounded by $\hat{Z}$ (see (3.18)).

Consequently, the process $\partial_x X$ does not reach zero almost surely. By continuity of $\partial_x X$ we observe

$$\partial_x X_\tau = \lim_{n\to\infty} \partial_x X_{\tau_n} > 0, \text{ a.s.}$$

This means that $\tau = T$ a.s. and $\hat{V}, \hat{Z}$ are defined on the whole interval $[t_0, T]$. Moreover, we have

$$\hat{V}_t = \frac{\partial_x P_t}{\partial_x X_t}, \; t \in [t_0, T], \text{ a.s.}$$

Thus, we obtain

$$\hat{V}_T = \frac{\partial_x P_T}{\partial_x X_T} = \frac{\xi'(X_T)\partial_x X_T}{\partial_x X_t} = \xi'(X_T), \text{ a.s.}$$

The definition of $\hat{V}$ in (3.12), written in the backward form, yields that $(\hat{V}, \hat{Z})$ solves the BSDE (3.11) on $[t_0, T]$. We have especially $\hat{V} \in \mathcal{S}^2_{t_0,T}$ because $\hat{V}$ is bounded, and $\hat{Z} \in \mathcal{H}^2_{t_0,T}$ because $\hat{Z}$ is a BMO-process on $[t_0, T]$ by the considerations above.

Note that $(\hat{V}, \hat{Z})$ is the only solution with bounded first component. Let $(\tilde{V}, \tilde{Z})$ be another solution to the BSDE (3.11) with bounded first component, i.e. there exists a constant $\tilde{K} \geq 0$ such that $|\tilde{V}_t| \leq \tilde{K}$ for all $t \in [t_0, T]$, $\mathbb{P}$-almost surely. We consider the generator $\tilde{\rho} : \Omega \times [t_0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by

$$\tilde{\rho}(t, v, z) := \rho(t, c(v), z),$$

where $c(v) = (-C \vee v) \wedge C$ and $C := K \vee \tilde{K}$. The function $\tilde{\rho}$ is defined for all $v \in \mathbb{R}$ in contrast to ρ and, in addition, Lipschitz continuous in v. Moreover, we observe that both $(\hat{V}, \hat{Z})$ and $(\tilde{V}, \tilde{Z})$ solve the BSDE with the standard parameters $(\xi'(X_T), -\tilde{\rho})$. But according to Theorem 2.3 the solution to this BSDE is unique and hence we obtain $(\tilde{V}, \tilde{Z}) = (\hat{V}, \hat{Z})$.

Finally, $(\hat{V}, \hat{Z})$ is even the unique solution in $\mathcal{S}^2_{t_0,T} \times \mathcal{H}^2_{t_0,T}$ to the BSDE with parameters $(\xi'(X_T), -\rho)$ because standard results about quadratic BSDEs (e.g. Theorem 2.6 in [8]) imply existence and uniqueness of a solution. □

In the following we consider $\hat{V}$, the time-continuous version of the gradient process, given by Lemma 3.12. A measure transformation simplifies the BSDE (3.11).

Lemma 3.13. *There exists a probability measure Q that is equivalent to $\mathbb{P}$, and a Q-Brownian motion W^Q such that $(\hat{V}, \hat{Z})$ is the unique solution to the BSDE*

$$\hat{V}_t = \xi'(X_T) + \int_t^T \varphi(s, \hat{V}_s)\,\mathrm{d}s - \int_t^T \hat{Z}_s\,\mathrm{d}W_s^Q, \tag{3.20}$$

for $t \in [t_0, T]$, where $\varphi : \Omega \times [t_0, T] \times \mathbb{R} \to \mathbb{R}$ is defined by $\varphi(s, v) := -B_s(1 - v)(v + 1)$.

Proof. Note that $\rho(t, \hat{V}_t, \hat{Z}_t) = -\varphi(t, \hat{V}_t) - \hat{Z}_t\psi(t, \hat{V}_t, \hat{Z}_t)$ for $(\omega, t) \in \Omega \times [t_0, T]$ with

$$\psi(\omega, t, v, z) := \frac{\Sigma_t(\omega)z}{1 - \Sigma_t(\omega)v}, \ (\omega, t, v, z) \in \Omega \times [t_0, T] \times [-K, K] \times \mathbb{R}.$$

We know that $\hat{V}$ and Σ are bounded by K and $L_{\sigma,z}$, respectively, and $|KL_{\sigma,z}| < 1$ (see Lemma 3.8 and Lemma 3.10). Therefore,

$$|\psi(t, \hat{V}_t, \hat{Z}_t)| \leq \frac{L_{\sigma,z}}{1 - KL_{\sigma,z}}|\hat{Z}_t|, \ t \in [t_0, T], \text{ a.s.}$$

This implies, in particular, that $\psi(\cdot, \hat{V}_\cdot, \hat{Z}_\cdot)$ is a BMO-process, since $\hat{Z}$ is a BMO-process and thus there exists a constant $C \geq 0$ such that for all $t \in [t_0, T]$

$$\mathbb{E}\left[\int_t^T \psi(s, \hat{V}_s, \hat{Z}_s)^2\,\mathrm{d}s \Big| \mathcal{F}_t\right] \leq \left(\frac{L_{\sigma,z}}{1 - KL_{\sigma,z}}\right)^2 \mathbb{E}\left[\int_t^T \hat{Z}_s^2\,\mathrm{d}s \Big| \mathcal{F}_t\right] \leq C, \text{ a.s.}$$

Applying Theorem A.1.2 in [6] implies that Q, defined by

$$Q(F) := \int_F \exp\left(\int_{t_0}^T \psi(s, \hat{V}_s, \hat{Z}_s)\,\mathrm{d}W_t - \frac{1}{2}\int_{t_0}^T \psi(s, \hat{V}_s, \hat{Z}_s)^2\,\mathrm{d}s\right)\,\mathrm{d}\mathbb{P}, \ F \in \mathcal{F},$$

is a probability measure that is equivalent to $\mathbb{P}$ with density

$$\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}} = \exp\left(\int_{t_0}^T \psi(s, \hat{V}_s, \hat{Z}_s)\,\mathrm{d}W_t - \frac{1}{2}\int_{t_0}^T \psi(s, \hat{V}_s, \hat{Z}_s)^2\,\mathrm{d}s\right).$$

If we define W^Q by

$$W_t^Q := W_t - \int_{t_0}^t \psi(s, \hat{V}_s, \hat{Z}_s)\,\mathrm{d}s, \ t \in [t_0, T],$$

Girsanov's theorem yields that W^Q is a Brownian motion w.r.t. Q. We note eventually that

$$\begin{aligned}
\hat{V}_t &= \xi'(X_T) - \int_t^T \rho(s, \hat{V}_s, \hat{Z}_s)\,\mathrm{d}s - \int_t^T \hat{Z}_s\,\mathrm{d}W_s \\
&= \xi'(X_T) + \int_t^T \varphi(s, \hat{V}_s)\,\mathrm{d}s + \int_t^T \hat{Z}_s\psi(s, \hat{V}_s, \hat{Z}_s)\,\mathrm{d}s - \int_t^T \hat{Z}_s\,\mathrm{d}W_s \\
&= \xi'(X_T) + \int_t^T \varphi(s, \hat{V}_s)\,\mathrm{d}s - \int_t^T \hat{Z}_s\,\mathrm{d}W_s^Q,
\end{aligned}$$

for all $t \in [t_0, T]$, $\mathbb{P}$-almost surely. □

Lemma 3.13 enables us to improve the bound given in Lemma 3.10. In particular, we can prove that the version $\hat{V}$ of the gradient process is uniformly bounded by a constant strictly smaller than 1, using repeatedly the comparison principle in a suitable way.

Lemma 3.14. *We have $|\hat{V}_t| \le q$ for all $t \in [t_0, T]$, $\mathbb{P}$-almost surely, where*

$$q := 1 - \left(1 - \frac{1}{1 + \gamma\delta_l}\right) e^{-2\|B\|_\infty T} < 1.$$

Proof. To determine the bound for $\hat{V}$ we consider the generator

$$\check{\varphi}(t, v) := \varphi(t, c(v)) = -B_t(1 - c(v))(1 + c(v)),\ (t, v) \in [t_0, T] \times \mathbb{R},$$

where $c(v) = (-1 \vee v) \wedge 1$. Because $(\xi'(X_T), \check{\varphi})$ are standard there exists a solution $(\check{V}, \check{Z})$ to the BSDE with parameters $(\xi'(X_T), \check{\varphi})$ according to Theorem 2.3. We are going to show that $\check{V}$ is bounded by q and then transfer this result to $\hat{V}$. Note that

$$0 \le \xi'(x) = \frac{1 - \gamma g''((\mathrm{Id} + \gamma g')^{-1}(x))}{1 + \gamma g''((\mathrm{Id} + \gamma g')^{-1}(x))} \le \frac{1}{1 + \gamma\delta_l},\ \omega \in \Omega,\ x \in \mathbb{R}, \tag{3.21}$$

because $g''(x) \ge \delta_l$. We have the following for $\check{V}$:

(1) $-1 \le \check{V}_t \le 1$ for all $t \in [t_0, T]$, $\mathbb{P}$-almost surely:

(i) The process $(V^{(1)}, Z^{(1)})$, given by $V_t^{(1)} := 1, Z_t^{(1)} := 0$, solves the BSDE with parameters $(1, \check{\varphi})$ on $[t_0, T]$ because $\check{\varphi}(t, 1) = 0$ for all $t \in [t_0, T]$, a.s. Using equation (3.21) we can apply the comparison principle (Theorem 2.5) to $(\check{V}, \check{Z})$ and $(V^{(1)}, Z^{(1)})$ and obtain that

$$\check{V}_t \le V_t^{(1)} = 1,\ t \in [t_0, T],\ \text{a.s.}$$

(ii) The process $(V^{(2)}, Z^{(2)})$, given by $V_t^{(2)} := -1, Z_t^{(2)} := 0$, solves the BSDE with parameters $(-1, \check{\varphi})$ on $[t_0, T]$ because $\check{\varphi}(t, -1) = 0$ for all $t \in [t_0, T]$, a.s. Using equation (3.21) we can apply the comparison principle (Theorem 2.5) to $(\check{V}, \check{Z})$ and $(V^{(2)}, Z^{(2)})$ and obtain that

$$\check{V}_t \ge V_t^{(2)} = -1,\ t \in [t_0, T],\ \text{a.s.}$$

(2) $-q \le \check{V}_t \le q$ for all $t \in [t_0, T]$, $\mathbb{P}$-almost surely: The property $\check{V}_t \in [-1, 1]$ implies that we can estimate the generator $\check{\varphi}$ in the following way:

$$\check{\varphi}(t, \check{V}_t) = -B(1 - \check{V}_t)(1 + \check{V}_t) \le 2\|B\|_\infty(1 - \check{V}_t), \tag{3.22}$$

$$\check{\varphi}(t, \check{V}_t) = -B(1 - \check{V}_t)(1 + \check{V}_t) \ge -2\|B\|_\infty(1 + \check{V}_t), \tag{3.23}$$

for all $t \in [t_0, T]$, a.s. Then we observe:

(i) The process $(V^{(3)}, Z^{(3)})$, given by $Z_t^{(3)} := 0$ and

$$V_t^{(3)} := 1 - \left(1 - \frac{1}{1 + \gamma\delta_l}\right) e^{-2\|B\|_\infty(T-t)},$$

solves the BSDE with parameters $\left(\frac{1}{1+\gamma\delta_l}, 2\|B\|_\infty(1-v)\right)$ on $[t_0,T]$, since

$$\begin{aligned}
&\frac{1}{1+\gamma\delta_l} + \int_t^T 2\|B\|_\infty \left(1 - V_s^{(3)}\right) \mathrm{d}s \\
&\qquad = \frac{1}{1+\gamma\delta_l} + \int_t^T 2\|B\|_\infty \left(1 - \frac{1}{1+\gamma\delta_l}\right) e^{-2\|B\|_\infty(T-s)} \,\mathrm{d}s \\
&\qquad = \frac{1}{1+\gamma\delta_l} + \left(1 - \frac{1}{1+\gamma\delta_l}\right)\left(1 - e^{-2\|B\|_\infty(T-t)}\right) = V_t^{(3)}, \;\; t \in [t_0,T].
\end{aligned}$$

Due to (3.21) and (3.22) we can apply the comparison principle to $(\check{V},\check{Z})$ and $(V^{(3)},Z^{(3)})$. This yields

$$\check{V}_t \le V_t^{(3)} \le 1 - \left(1 - \frac{1}{1+\gamma\delta_l}\right) e^{-2\|B\|_\infty T} = q, \;\; t \in [t_0,T], \;\; \text{a.s.}$$

(ii) The process $(V^{(4)},Z^{(4)})$, given by $V_t^{(4)} := e^{-2\|B\|_\infty(T-t)} - 1$ and $Z_t^{(4)} := 0$, solves the BSDE with parameters $(0, -2\|B\|_\infty(1+v))$ on $[t_0,T]$, since

$$\begin{aligned}
-\int_t^T 2\|B\|_\infty \left(1 + V_s^{(4)}\right) \mathrm{d}s &= -\int_t^T 2\|B\|_\infty e^{-2\|B\|_\infty(T-s)} \,\mathrm{d}s \\
&= -1 + e^{-2\|B\|_\infty(T-t)} = V_t^{(4)}, \;\; t \in [t_0,T].
\end{aligned}$$

Again the comparison principle implies

$$\check{V}_t \ge V_t^{(4)} \ge e^{-2\|B\|_\infty T} - 1 \ge \left(1 - \frac{1}{1+\gamma\delta_l}\right) e^{-2\|B\|_\infty T} - 1 = -q, \;\; t \in [t_0,T], \;\; \text{a.s.},$$

because of (3.21) and (3.23).

Note that the generator φ of the BSDE (3.20) might not be Lipschitz continuous. But we can alter φ to gain a Lipschitz continuous generator $\tilde{\varphi}$, defined by

$$\tilde{\varphi}(t,v) := \varphi(t, (v \vee -\tilde{K}) \wedge \tilde{K}), \;\; (t,v) \in [t_0,T] \times \mathbb{R},$$

for $\tilde{K} := K \vee 1$. As a matter of fact, $(\hat{V},\hat{Z})$ solves the modified BSDE with parameters $(\xi',\tilde{\varphi})$ because $\hat{V}$ is bounded by K (see the construction of $\hat{V}$ in the proof of Lemma 3.12). Theorem 2.3 implies that it is the unique solution to that BSDE. But $(\check{V},\check{Z})$ also solves the BSDE with parameters $(\xi',\tilde{\varphi})$ because $|\check{V}_t| \le 1 \le \tilde{K}$, $t \in [t_0,T]$, a.s., and hence $\check{\varphi}(t,\check{V}_t) = \tilde{\varphi}(t,\check{V}_t)$ for all $t \in [t_0,T]$, a.s. Uniqueness of the solution implies $(\hat{V},\hat{Z}) = (\check{V},\check{Z})$ and thus also $\hat{V}$ is bounded by the constant q. □

Now we sum up all the statements of this section. In particular, we transfer the bound of $\hat{V}$ to the gradient process V.

Proposition 3.15. *The following holds true:*

1. *For all $t \in I_{max}$ we have that $|u_x(t,x)| \leq q$ for $\mathbb{P} \otimes \lambda$-almost all $(\omega, x) \in \Omega \times \mathbb{R}$. Consequently, $L_{u,x} \leq q$ and $I_{max} = [0,T]$.*

2. *There exists a unique strongly regular, progressively measurable and Lipschitz continuous decoupling field u for the auxiliary FBSDE* (3.4) *on $[0,T]$ such that*

$$|u_x(t,x)| \leq q$$

for all $t \in [0,T]$ and for $\mathbb{P} \otimes \lambda$-almost all $(\omega, x) \in \Omega \times \mathbb{R}$. In particular, we have $L_{u,x} \leq q$.

3. *For $x \in \mathbb{R}$ there exists a unique solution $(X^x, P^x, \tilde{Z}^x)$ to the auxiliary FBSDE* (3.4).

Proof. The version $\hat{V}$ of the gradient process V is bounded by q due to Lemma 3.14. The construction of $\hat{V}$ especially implies that $\hat{V}_{t_0} = V_{t_0}$ a.s. and thus

$$|u(t_0, x)| = |u(t_0, X_{t_0})| = |V_{t_0}| = |\hat{V}_{t_0}| \leq q, \text{ a.s.}$$

The choice of x on page 47 was outside of a Lebesgue null set. Therefore, $u(\cdot, t_0, \cdot)$ is essentially bounded by q. This bound does not depend on the choice of t_0. Because we have arbitrarily selected $t_0 \in I_{max}$ on page 44, we obtain for all $t \in I_{max}$

$$|u_x(\omega, t, x)| \leq q, \tag{3.24}$$

for $\mathbb{P} \otimes \lambda$-almost all $(\omega, x) \in \Omega \times \mathbb{R}$. Let now $t \in I_{max}$ be fixed for the following considerations. We have according to Lemma A.2.1 in [6] that for almost all $x_1, x_2 \in \mathbb{R}, x_1 \leq x_2$

$$\int_{x_1}^{x_2} u_x(t,y)\,\mathrm{d}y = u(t,x_2) - u(t,x_1), \text{ a.s.,}$$

and hence by (3.24)

$$|u(t,x_2) - u(t,x_1)| \leq \int_{x_1}^{x_2} |u_x(t,y)|\,\mathrm{d}y \leq q|x_2 - x_1|, \text{ a.s.}$$

Now Lemma A.2.2 in [6] yields that we have for the Lipschitz constant $L_{u(t,\cdot),x}$ of $u(t,\cdot)$

$$L_{u(t,\cdot),x} \leq q,$$

because $u(\omega, t, \cdot)$ is Lipschitz continuous for all $(\omega, t) \in \Omega \times [t_0, T]$. Since the choice of $t \in I_{max} = (t_{min}, T]$ was arbitrary, we see that

$$\lim_{t \downarrow t_{min}} L_{u(t,\cdot),x} \leq L_{u,x} \leq q < 1 \leq L_{\sigma,z}^{-1},$$

which contradicts Proposition 2.22. Finally, Theorem 2.21 implies that:

1. $I_{max} = [0,T]$.

2. There exists a unique strongly regular decoupling field u on $[0,T]$ that is progressively measurable and Lipschitz continuous (cf. Remark 2.14). The results in this section are especially true for $I_{max} = [0,T]$ and thus $|u_x(\cdot, t, \cdot)|$ is essentially bounded by q for all $t \in [0,T]$ and $L_{u,x} \leq q$.

3. There exists a unique solution $(X^x, P^x, \tilde{Z}^x)$ to the FBSDE (3.4) for every $x \in \mathbb{R}$.

This concludes the proof. □

3.5 The results

In sections 3.3 and 3.4 we have shown that the auxiliary FBSDE (3.4) has a solution and that we can transfer this solvability to our original FBSDE (3.3). Furthermore, we have proven in Proposition 3.4 with the help of the maximum principle that there is an optimal control. We now state all the main results of this chapter.

Theorem 3.16. *We have that:*

1. *For all $m \in \mathbb{R}$ there exists a solution (M^m, Y^m, Z^m) to the FBSDE* (3.3)*.*
2. *The function $\nu : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$, defined by*

$$\nu(\omega, t, m) := \frac{1}{2\gamma} \left(Id - u(\omega, t, \cdot) \right) \left(\left(Id + u(\omega, t, \cdot) \right)^{-1} (2m) \right),$$

is a decoupling field for the FBSDE (3.3)*, i.e. $Y_t^m = \nu(t, M_t^m)$ a.s. for all $t \in [0,T]$.*

3. *An optimal control for our problem (P) is given by $\hat{\alpha}$, defined as*

$$\hat{\alpha}_t := a_*(t, -Z_t^m) = \begin{cases} l & , -Z_t^m < f_a(t,l), \\ f_a^{-1}(t, -Z_t^m) & , -Z_t^m \in [f_a(t,l), f_a(t,r)], \\ r & , -Z_t^m > f_a(t,r). \end{cases}$$

Proof. 1. Let $m \in \mathbb{R}$. At first, we show how we can calculate the initial value for the auxiliary FBSDE (3.4). To that end, we consider the decoupling field u of Proposition 3.15. We have, in particular, that $u(\omega, t, \cdot)$ is Lipschitz continuous for all $(\omega, t) \in \Omega \times [0,T]$. Moreover, the progressive measurability of $u(\cdot, \cdot, x)$ for all fixed $x \in \mathbb{R}$ implies that $u(\cdot, 0, x)$ is almost surely constant.

Now we consider the continuous mapping h, defined by

$$h(\omega, t, x) := (x + u(\omega, t, x)), \ (\omega, t, x) \in \Omega \times [0,T] \times \mathbb{R}.$$

Let $x_1, x_2 \in \mathbb{R}, x_1 < x_2$. Using the Lipschitz continuity of u with Lipschitz constant $L_{u,x} \leq q$, we have for all $(\omega, t) \in \Omega \times [0,T]$

$$\begin{aligned} h(t, x_2) - h(t, x_1) &= (x_2 - x_1) + (u(t, x_2) - u(t, x_1)) \\ &\geq (x_2 - x_1) - q(x_2 - x_1) = (1-q)(x_2 - x_1) > 0, \end{aligned}$$

since $q < 1$ and $x_2 > x_1$. This yields that h is strictly increasing. Moreover, the estimate implies that $\lim_{x \to \pm\infty} h(t,x) = \pm\infty$. Consequently, the mapping h is a bijection from $\mathbb{R}$ onto $\mathbb{R}$ and thus the inverse in x exists. We observe that for all $x \in \mathbb{R}$ the function $h(\cdot, 0, x)$ is constant almost surely and thus also $h^{-1}(\cdot, 0, x)$ is constant almost surely.

Let now $x = h^{-1}(0, 2m)$, i.e. $x \in \mathbb{R}$ is chosen such that $x = h^{-1}(0, 2m)$ a.s. Proposition 3.15 implies that there exists a solution $(X^x, P^x, \tilde{Z}^x)$ on $[0,T]$ with initial value x. By Proposition 3.7 we obtain that (M^m, Y^m, Z^m), defined by

$$M_s^m := \frac{X_s^x + P_s^x}{2}, \qquad Y_s^m := \frac{X_s^x - P_s^x}{2\gamma}, \qquad Z_s^m := -\tilde{a}_*^{-1}\left(s, \tilde{Z}_s^x\right), \ s \in [0,T],$$

solves the FBSDE (3.2) on $[0,T]$ with initial value m. The initial value is indeed equal to m since

$$M_0^m = \frac{1}{2}\left(X_0^x + P_0^x\right) = \frac{1}{2}\left(X_0^x + u(0, X_0^x)\right) = \frac{1}{2}h(0,x) = h(0, h^{-1}(0, 2m)) = m, \text{ a.s.}$$

This calculation motivated the definition of h.

2. The function ν is well-defined because the mapping h is invertible for all fixed $(\omega, t) \in \Omega \times [0,T]$ as we have proven in the first part. We check if ν is indeed a decoupling field for the FBSDE (3.3). Note that $u(T,\cdot) = \xi$ a.s. and therefore by the definition of ξ in (3.5)

$$\left(\mathrm{Id} + u(T,\cdot)\right)(x) = x + \xi(x) = x + 2(\mathrm{Id} + \gamma g')^{-1}(2x) - x = 2(\mathrm{Id} + \gamma g')^{-1}(x), \text{ a.s.,}$$

for all $x \in \mathbb{R}$. This implies that $(\mathrm{Id} + u(T,\cdot))^{-1}(x) = (\mathrm{Id} + \gamma g')\left(\frac{x}{2}\right)$ a.s. for all $x \in \mathbb{R}$. Using again the definition of ξ we have that

$$\begin{aligned}
\nu(T,x) &= \frac{1}{2\gamma}\left(\mathrm{Id} - u(T,\cdot)\right)\left(\left(\mathrm{Id} + u(T,\cdot)\right)^{-1}(2x)\right) \\
&= \frac{1}{2\gamma}\left(\mathrm{Id} - \xi\right)\left((\mathrm{Id} + \gamma g')(x)\right) \\
&= \frac{1}{2\gamma}\left[x + \gamma g'(x) - \xi\left(\left(\mathrm{Id} + \gamma g'\right)(x)\right)\right] \\
&= \frac{1}{2\gamma}\left(x + \gamma g'(x) - 2x + (\mathrm{Id} + \gamma g')(x)\right) = g'(x), \text{ a.s.}
\end{aligned}$$

Hence (1) of Definition 2.9 is fulfilled. Moreover, for all $t_1, t_2 \in [0,T], t_1 \leq t_2$ and any $\mathcal{F}_{t_1}$-measurable M_{t_1} we can find processes M, Y, Z satisfying properties (2)-(5) of Definition 2.9 as one can see as follows. Because u is a decoupling field for the auxiliary FBSDE (3.4), there are processes $X, P, \tilde{Z}$ fulfilling the properties (2)-(5) of Definition 2.9 for the initial condition

$$X_{t_1} := M_{t_1} + \gamma \nu(t_1, M_{t_1})$$

and the auxiliary FBSDE (3.4). If one defines M, Y, Z by

$$M_s := \frac{1}{2}(X_s + P_s), \; Y_s := \frac{1}{2\gamma}(X_s + P_s), \; Z_s := -\tilde{a}_*^{-1}(s, \tilde{Z}_s), \; s \in [t_1, t_2],$$

one can show in the same manner as in Proposition 3.7 that M, Y, Z have the properties (2)-(4). We show that the decoupling condition (5) is satisfied on $[t_1, t_2]$. Note that for all $t \in [t_1, t_2]$

$$2M_t = X_t + P_t = \left(\mathrm{Id} + u(t,\cdot)\right)(X_t), \text{ a.s.,}$$

and therefore

$$\begin{aligned}
\nu(t, M_t) &= \frac{1}{2\gamma}\left(\mathrm{Id} - u(t,\cdot)\right)\left(\left(\mathrm{Id} + u(t,\cdot)\right)^{-1}(2M_t)\right) \\
&= \frac{1}{2\gamma}\left(\mathrm{Id} - u(t,\cdot)\right)(X_t) \\
&= \frac{1}{2\gamma}\left(X_t - u(t, X_t)\right) = \frac{1}{2\gamma}\left(X_t - P_t\right) = Y_t, \text{ a.s.}
\end{aligned}$$

3. Proposition 3.4 implies that $\hat{\alpha}$ is indeed an optimal control to problem (P) since there exists a solution (M^m, Y^m, Z^m) to the FBSDE (3.3) for every initial value $m \in \mathbb{R}$. □

Chapter 4

Connecting decoupling field and value function

In this chapter we relate the FBSDE approach via the maximum principle and the HJB approach. In particular, we show that the weak derivative of the value function is equal to the decoupling field for the optimal control problem presented in [1]. This control problem is very similar to the one that we consider in this thesis. But the control processes can take values in $\mathbb{R}$ and the function f might depend on the state of our underlying process in contrast to our problem in chapter 3. The results, however, are very much alike.

This chapter is organized as follows: At first, we introduce the control problem and summarize all important results to make this chapter self-contained. In the second part, we present in an example that the aforementioned relation is true. We integrate the decoupling field and choose the constant of integration such that we obtain the value function. Finally, the last part presents the link between the value function and the decoupling field in the general setting. We calculate the weak derivative of the value function and show that it is equal to the decoupling field.

4.1 Problem formulation and important results

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and let W be a Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $T > 0$ be a finite time horizon. Define furthermore the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ by $\mathcal{F}_t := \sigma(\mathcal{N}, (W_s)_{s\in[0,t]})$ for $t \in [0,T]$ with $\mathcal{N}$ being the set of all $\mathbb{P}$-null sets in $\mathcal{F}$.

For $t \in [0,T]$ we denote by $\mathcal{A}(t)$ the set of all admissible controls defined on $[t,T]$, i.e. $\mathcal{A}(t)$ contains the progressively measurable processes $\alpha : \Omega \times [t,T] \to \mathbb{R}$ such that $\mathbb{E}\int_t^T \alpha_s^2 \,\mathrm{d}s < \infty$. Moreover, let $b, B : \Omega \times [0,T] \to \mathbb{R}$ be progressively measurable and bounded processes. We consider the stochastic process $M^{t,m,\alpha}$ driven by the stochastic differential equation

$$M_s^{t,m,\alpha} = m + \int_t^s (b_r + B_r M_r^{t,m,\alpha}) \,\mathrm{d}r + \int_t^s \alpha_r \,\mathrm{d}W_r, \; s \in [t,T], \tag{4.1}$$

which is controlled by α and starts in $m \in \mathbb{R}$ at time $t \in [0,T]$. In order to introduce the value function, we define the measurable mappings $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $g : \Omega \times \mathbb{R} \to \mathbb{R}$, which can be interpreted as running and terminal costs, respectively. We require that the mapping $(\omega, t) \mapsto f(\omega, t, m, a)$ is progressively measurable for all

$(m,a) \in \mathbb{R}^2$. Moreover, we assume that f and g satisfy the conditions (C1)-(C3) adapted to this setting (cf. p. 33), i.e. we assume:

(D1) For all fixed $(\omega,t) \in \Omega \times [0,T]$ the mapping f is convex in m and strictly convex in a, and g is convex in m. In addition, it holds

$$\|f(\cdot,\cdot,0,0)\|_\infty + \|g(\cdot,0)\|_\infty < \infty,$$

where $\|h\|_\infty$ denotes the essential supremum of $|h|$ for some measurable function h.

(D2) The functions $g(\omega,\cdot)$ and $f(\omega,t,\cdot,\cdot)$ are two times continuously differentiable for all $(\omega,t) \in \Omega \times [0,T]$. Additionally, the derivatives g', f_a, f_m are Lipschitz continuous in the variables m and a, i.e. there exists a constant $L \geq 0$ such that for all $\omega \in \Omega$, $t \in [0,T]$, $m_1, m_2 \in \mathbb{R}$, $a_1, a_2 \in \mathbb{R}$

$$\begin{aligned}
|g'(\omega,m_1) - g'(\omega,m_2)| &\leq L|m_1 - m_2|, \\
|f_m(\omega,t,m_1,a_1) - f_m(\omega,t,m_2,a_2)| &\leq L\left(|m_1-m_2| + |a_1-a_2|\right), \\
|f_a(\omega,t,m_1,a_1) - f_a(\omega,t,m_2,a_2)| &\leq L\left(|m_1-m_2| + |a_1-a_2|\right).
\end{aligned}$$

Furthermore the following property holds:

$$\|f_m(\cdot,\cdot,0,0)\|_\infty + \|f_a(\cdot,\cdot,0,0)\|_\infty + \|g'(\cdot,0)\|_\infty < \infty.$$

(D3) There is a constant $\delta_l > 0$ such that

$$f_{aa}(\omega,t,m,a) \geq \delta_l$$

for all $(\omega,t,m,a) \in \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}$.

Our goal is to minimize the target function

$$J(t,m,\alpha) := \mathbb{E}\left[\int_t^T f\left(s, M_s^{t,m,\alpha}, \alpha_s\right) \, \mathrm{d}s + g\left(M_T^{t,m,\alpha}\right) \middle| \mathcal{F}_t\right],$$

$(t,m) \in [0,T] \times \mathbb{R}, \alpha \in \mathcal{A}(t)$, over all admissible controls. To that end, we introduce the value function $v : \Omega \times [0,T] \times \mathbb{R} \to [-\infty,\infty]$, defined by

$$v(t,m) := \operatorname*{ess\,inf}_{\alpha \in \mathcal{A}(t)} J(t,m,\alpha), \ (t,m) \in [0,T] \times \mathbb{R}.$$

Note that both value and target function depend on ω, and that more details about this definition can be found in section 2.4. In order to obtain an explicit representation of v, one has to find for $(t,m) \in [0,T] \times \mathbb{R}$ an optimal control $\alpha \in \mathcal{A}$, i.e. a control α that fulfils

$$v(t,m) = \mathbb{E}\left[\int_t^T f\left(s, M_s^{t,m,\alpha}, \alpha_s\right) \, \mathrm{d}s + g\left(M_T^{t,m,\alpha}\right) \middle| \mathcal{F}_t\right], \text{ a.s.}$$

Therefore, we apply the maximum principle in the general form (Theorem 2.28), for which we have to study the solvability of the adjoint FBSDE

$$\begin{aligned}
M_s^{t,m} &= m + \int_t^s (b_r + B_r M_r^{t,m}) \, \mathrm{d}r + \int_t^s f_a^{-1}(r, M_r^{t,m}, -Z_r^{t,m}) \, \mathrm{d}W_r, \\
Y_s^{t,m} &= g'(M_T^{t,m}) - \int_s^T Z_r^{t,m} \, \mathrm{d}W_r \\
&\quad + \int_s^T \left(B_r Y_r^{t,m} + f_m\left(r, M_r^{t,m}, f_a^{-1}(r, M_r^{t,m}, -Z_r^{t,m})\right)\right) \mathrm{d}r, \ s \in [t,T],
\end{aligned} \tag{4.2}$$

for given $(t,m) \in [0,T]\times\mathbb{R}$, because the function $f_a^{-1}(s,m,-z)$ minimizes the Hamiltonian (cf. [1, p. 4]). The above FBSDE is solved in [1] with the help of an auxiliary FBSDE and the method of decoupling fields like in chapter 3. In particular, one has the following results.

Proposition 4.1. *We have:*

1. *There exists a Lipschitz continuous and progressively measurable decoupling field u on $[0,T]$ for the FBSDE* (4.2)*.*

2. *For all $(t,m) \in [0,T]\times\mathbb{R}$ there exists a solution $(M^{t,m}, Y^{t,m}, Z^{t,m}) \in \mathcal{S}^2_{t,T}\times\mathcal{S}^2_{t,T}\times\mathcal{H}^2_{t,T}$ to the FBSDE* (4.2)*.*

3. *For all $s \in [t,T]$ we have that $M^{t,m}_s, Y^{t,m}_s$ are weakly differentiable w.r.t. m, and $Z^{t,m}$ is weakly differentiable w.r.t. m. It holds that*

$$\begin{aligned} &\operatorname*{ess\,sup}_{m\in\mathbb{R}} \sup_{s\in[t,T]} \mathbb{E}\left[\left(\partial_m M^{t,m}_s\right)^2\right] < \infty, \\ &\operatorname*{ess\,sup}_{m\in\mathbb{R}} \sup_{s\in[t,T]} \mathbb{E}\left[\left(\partial_m Y^{t,m}_s\right)^2\right] < \infty, \\ &\operatorname*{ess\,sup}_{m\in\mathbb{R}} \mathbb{E}\left[\int_t^T \left(\partial_m Z^{t,m}_s\right)^2 \,\mathrm{d}s\right] < \infty. \end{aligned} \tag{4.3}$$

   Moreover, $\partial_m M^{t,m}, \partial_m Y^{t,m} \in \mathcal{S}^2_{t,T}$ and $\partial_m Z^{t,m} \in \mathcal{H}^2_{t,T}$ for almost every $m \in \mathbb{R}$.

4. *For all $(t,m) \in [0,T]\times\mathbb{R}$ the control $\hat\alpha \in \mathcal{A}(t)$, defined by $\hat\alpha_s := f_a^{-1}(s, M^{t,m}_s, -Z^{t,m}_s)$ for $s \in [t,T]$, is an optimal control, i.e. $v(t,m) = J(t,m,\alpha)$ a.s.*

Sketch of the proof. According to Proposition 4.5 and Theorem 4.6 in [1], the auxiliary FBSDE of (4.2) has a strongly regular decoupling field $\tilde u$ that satisfies:

- $\tilde u(\omega,t,\cdot) : \mathbb{R}\to\mathbb{R}$ is Lipschitz continuous and invertible a.s. for all $t \in [0,T]$,
- $\tilde u^{-1}(\omega,t,\cdot) : \mathbb{R}\to\mathbb{R}$ is Lipschitz continuous a.s. for all $t \in [0,T]$,
- for the weak derivative of $\tilde u$ we have $q \le \partial_x \tilde u(t,x) \le 1$ for all $t \in [0,T]$ and almost all $(\omega,x) \in \Omega\times\mathbb{R}$, where $q \in (0,1)$ is some constant.

In particular, the strong regularity of $\tilde u$ implies that there exists a solution $(X^{t,x}, P^{t,x}, \tilde Z^{t,x})$ to the auxiliary FBSDE for all $(t,x) \in [0,T]\times\mathbb{R}$. The random variables $X^{t,x}_s, P^{t,x}_s$ and the process $\tilde Z^{t,x}$ are weakly differentiable w.r.t. x for all $s \in [t,T]$ and the property (2.7) is fulfilled. Now we show the statements of the proposition:

1. Follows directly from Theorem 4.6 in [1].

2. Let $(t,m) \in [0,T]\times\mathbb{R}$. We construct the solution $(M^{t,m}, Y^{t,m}, Z^{t,m})$ as follows. The property of $\tilde u$ being a decoupling fields implies that there exists a solution

$$(X,P,\tilde Z) = \left(X^{t,\tilde u^{-1}(t,m)}, P^{t,\tilde u^{-1}(t,m)}, \tilde Z^{t,\tilde u^{-1}(t,m)}\right)$$

to the auxiliary FBSDE with the $\mathcal{F}_t$-measurable initial condition $\tilde{u}^{-1}(t,m)$. One can now recover $(M^{t,m}, Y^{t,m}, Z^{t,m})$ by the transformation

$$\begin{aligned} M_s^{t,m} &:= P_s^{t,\tilde{u}^{-1}(t,m)}, \qquad Y_s^{t,m} := C\left(X_s^{t,\tilde{u}^{-1}(t,m)} - P_s^{t,\tilde{u}^{-1}(t,m)}\right), \\ Z_s^{t,m} &:= -f_a\left(s, P_s^{t,\tilde{u}^{-1}(t,m)}, \tilde{Z}_s^{t,\tilde{u}^{-1}(t,m)}\right), \ s \in [t,T], \end{aligned} \tag{4.4}$$

where $C > 0$ is some constant (cf. Lemma 4.1 in [1]). As in Proposition 3.7 one can show that $M^{t,m}, Y^{t,m} \in \mathcal{S}^2_{t,T}$ and $Z^{t,m} \in \mathcal{H}^2_{t,T}$.

3. Let $t \in [0,T]$. The solution $(X, P, \tilde{Z})$ to the auxiliary FBSDE is weakly differentiable w.r.t. the initial value. As aforementioned, the transformation in (4.4) and the chain rule in Proposition 2.40 imply that also $M_s^{t,m}$ and $Y_s^{t,m}$ are weakly differentiable w.r.t. m for all $s \in [t,T]$, and the weak derivative of $M^{t,m}$ is, for instance, given by

$$\partial_m M_s^{t,m} = \frac{\partial_x P_s^{t,\tilde{u}^{-1}(t,m)}}{\tilde{u}'(t, \tilde{u}^{-1}(t,m))}, \text{ a.s.,} \tag{4.5}$$

for almost every $m \in \mathbb{R}$. Moreover, the process $Z^{t,m}$ is weakly differentiable according to the chain rules in Proposition 2.38 and Proposition 2.40.

We now show that the property in (4.3) holds true for $\partial_m M^{t,m}$. The proof for the processes $\partial_m Y^{t,m}, \partial_m Z^{t,m}$ is analogue.

Without loss of generality we can assume that $\Omega = \Omega_1 \times \Omega_2$, and that the projections π_1, π_2 are independent such that $\mathcal{F}_t = \sigma(\pi_1) \vee \mathcal{N}$ and $\sigma(\pi_2) = \sigma\big((W_r - W_t)_{r\in[t,T]}\big)$. This means that we decompose Ω into the past and the future at time t. Then the probability measure $\mathbb{P}$ is a product measure, i.e. $\mathbb{P} = \mathbb{P}_1 \otimes \mathbb{P}_2$ for $\mathbb{P}_i := \mathbb{P} \circ \pi_i^{-1}, i = 1,2$.

Note that now for all $m \in \mathbb{R}$ and almost all $\omega_1 \in \Omega_1$ we have that $\tilde{u}^{-1}((\omega_1,\cdot),t,m)$ is constant since $\tilde{u}^{-1}(\cdot,t,m)$ is $\mathcal{F}_t$-measurable. Then for all $m \in \mathbb{R}$ and almost all $\omega_1 \in \Omega_1$ fixed there exists a solution

$$\left(X^{t,\tilde{u}^{-1}((\omega_1,\cdot),t,m)}(\omega_1,\cdot), P^{t,\tilde{u}^{-1}((\omega_1,\cdot),t,m)}(\omega_1,\cdot), \tilde{Z}^{t,\tilde{u}^{-1}((\omega_1,\cdot),t,m)}(\omega_1,\cdot)\right)$$

to a modified auxiliary FBSDE on $[t,T]$, where all parameters only depend on ω_2 since ω_1 is fixed. These parameters are still progressively measurable and satisfy SLC. Hence the solution is weakly differentiable and fulfils (2.7). This means especially that

$$\operatorname*{ess\,sup}_{(\omega_1,m)} \sup_{s\in[t,T]} \operatorname{ess\,sup} \mathbb{E}\left[\left(\partial_x P_s^{t,\tilde{u}^{-1}((\omega_1,\cdot),t,m)}(\omega_1,\cdot)\right)^2 \Big| \mathcal{F}_t\right] < \infty,$$

or equivalently, there exists a constant $C > 0$ such that for almost all $\omega_1 \in \Omega_1$ and almost every $m \in \mathbb{R}$

$$\mathbb{E}\left[\left(\partial_x P_s^{t,\tilde{u}^{-1}((\omega_1,\cdot),t,m)}(\omega_1,\cdot)\right)^2\right] \le C,$$

for all $s \in [t,T]$. Equation (4.5) and the estimate on the weak derivative of $\tilde{u}$ imply

$$\mathbb{E}\left[\left(\partial_m M_s^{t,m}(\omega_1,\cdot)\right)^2\right] \le \frac{1}{q^2}\mathbb{E}\left[\left(\partial_x P_s^{t,\tilde{u}^{-1}((\omega_1,\cdot),t,m)}(\omega_1,\cdot)\right)^2\right] \le \frac{C}{q^2},$$

for all $s \in [t,T]$, almost all $\omega_1 \in \Omega_1$ and almost every $m \in \mathbb{R}$. By calculating the expectation over $\omega = (\omega_1, \omega_2)$ we obtain using the above estimate and Fubini's theorem

$$\begin{aligned}
\mathbb{E}\left[\left(\partial_m M_s^{t,m}\right)^2\right] &= \int_\Omega \left(\partial_m M_s^{t,m}(\omega_1,\omega_2)\right)^2 \, \mathrm{d}\mathbb{P}(\omega_1,\omega_2) \\
&= \int_{\Omega_1}\int_{\Omega_2} \left(\partial_m M_s^{t,m}(\omega_1,\omega_2)\right)^2 \, \mathrm{d}\mathbb{P}_2(\omega_2)\, \mathrm{d}\mathbb{P}_1(\omega_1) \\
&= \int_{\Omega_1} \mathbb{E}\left[\left(\partial_m M_s^{t,m}(\omega_1,\cdot)\right)^2\right] \, \mathrm{d}\mathbb{P}_1(\omega_1) \\
&\le \int_{\Omega_1} \frac{1}{q^2}\mathbb{E}\left[\left(\partial_x P_s^{t,\tilde{u}^{-1}((\omega_1,\cdot),t,m)}(\omega_1,\cdot)\right)^2\right] \, \mathrm{d}\mathbb{P}_1(\omega_1) \le \frac{C}{q^2},
\end{aligned}$$

for all $s \in [t,T]$ and almost all $m \in \mathbb{R}$. Finally,

$$\operatorname*{ess\,sup}_{m\in\mathbb{R}} \sup_{s\in[t,T]} \mathbb{E}\left[\left(\partial_m M_s^{t,m}\right)^2\right] < \infty.$$

Furthermore, one can show as in Theorem 2.21 that $\partial_m M^{t,m}, \partial_m Y^{t,m} \in \mathcal{S}^2_{t,T}$, using the dynamics of these processes that can be calculated like in Proposition 3.8. The property $\partial_m Z^{t,m} \in \mathcal{H}^2_{t,T}$ follows from (4.3)

4. Let $(t,m) \in [0,T] \times \mathbb{R}$. We prove that $\hat{\alpha} \in \mathcal{A}(t)$ in Lemma 4.5. Note that our assumptions and the existence of a solution to (4.2) ensure that the requirements of the maximum principle (Theorem 2.28) are satisfied for the control α. Hence we obtain that α is optimal and $v(t,m) = J(t,m,\alpha)$ a.s.

□

4.2 An example

In this section we consider an example presented in [1, p. 15ff], where one can explicitly calculate the decoupling field u. Integration and the choice of the right constant of integration shows that the antiderivative of the decoupling field u equals the value function v, or in other words, the derivative of v equals u.

Let $\bar{a} : [0,T] \to \mathbb{R}$ be a bounded and continuous deterministic process. We define the constants $l > 0, L \ge 0, b = 0$ and $B \in \mathbb{R}$, and suppose that $f(t,m,a) := lm^2 + (a - \bar{a}_t)^2$ and $g(m) := Lm^2$. Note that we are in a Markovian framework and that equation (4.1) takes the form

$$M_s^{t,m,\alpha} = m + \int_t^s B M_r^{t,m,\alpha}\,\mathrm{d}r + \int_t^s \alpha_r\,\mathrm{d}W_r,\ s \in [t,T],\ (t,m) \in [0,T]\times\mathbb{R}. \tag{4.6}$$

Here we interpret $M^{t,m,\alpha}$ as the position of a particle in a medium with temperature process α. The natural temperature that surrounds our medium is given by the process $\bar{a}$. Then the function f describes the costs that any cooling or heating of the medium causes. Any deviation of the medium's temperature from the natural temperature generates quadratic costs. Moreover, we interpret the function g as quadratic penalization if a control does not steer the particle close enough to zero. The optimal control problem is characterized by the goal of steering the particle as close as possible to zero, while reducing the costs of temperature changes.

The adjoint FBSDE (4.2) is solved in [1] with the help of an auxiliary FBSDE just as in chapter 3. In particular, one studies the dynamics of the gradient process V, which are in this example given by the Riccati equation

$$V_t' = 2BV_t + (l - 2B)V_t^2, \qquad V_T = \frac{1}{1+L}. \tag{4.7}$$

Depending on B we obtain the following solution for V

$$V_t = \begin{cases} \dfrac{1}{l(T-t)+1+L} & , B = 0, \\[2ex] \dfrac{e^{-2B(T-t)}}{1+L+\frac{l-2B}{2B}\left(1-e^{-2B(T-t)}\right)} & , B \neq 0. \end{cases}$$

Consequently, the decoupling field $\tilde{u}$ for the auxiliary FBSDE is given by $\tilde{u}(t,x) := V_t x$. Theorem 4.6 in [1] yields that

$$u(t,m) := 2m\left(\frac{1}{V_t} - 1\right), \ (t,m) \in [0,T] \times \mathbb{R}$$

is a decoupling field u to the FBSDE (4.2). We verify this claim by calculating the decoupling condition in equation (4.8) below.

Note that we can recover a solution $(M^{t,m}, Y^{t,m}, Z^{t,m})$ to the FBSDE (4.2) from the auxiliary FBSDE like in the proof of Proposition 4.1. In more detail, let $t \in [0,T]$ be the initial time, and let $m \in \mathbb{R}$, $x := V_t^{-1}m$ be the initial values. Then there exists an explicit solution $(X^{t,x}, P^{t,x}, \tilde{Z}^{t,x})$ to the auxiliary FBSDE with initial value x (see Proposition 5.1 in [1]). We can now recover a solution $(M^{t,m}, Y^{t,m}, Z^{t,m})$ to the FBSDE (4.2) by

$$M_s^{t,m} := P_s^{t,x}, \quad Y_s^{t,m} := 2\left(X_s^{t,x} - P_s^{t,x}\right), \quad Z_s^{t,m} := -f_a(s, P_s^{t,x}, \tilde{Z}_s^{t,x}) = -2(\tilde{Z}_s^{t,x} - \bar{a}_s),$$

for $s \in [0,T]$. These considerations entail that the function u defined above is indeed a decoupling field since for all $(t,m) \in [0,T] \times \mathbb{R}$ and $s \in [t,T]$

$$\begin{aligned} u(s, M_s^{t,m}) &= 2M_s^{t,m}\left(\frac{1}{V_s} - 1\right) = 2P_s^{t,x}\left(\frac{1}{V_s} - 1\right) = \frac{2}{V_s}\tilde{u}\left(s, X_s^{t,x}\right) - 2P_s^{t,x} \\ &= 2X_s^{t,x} - 2P_s^{t,x} = Y_s^{t,m}, \ \text{a.s.}, \end{aligned} \tag{4.8}$$

where $x := V_t^{-1}m$. In the calculation we have used the connection between the processes $(M^{t,m}, Y^{t,m}, Z^{t,m})$ and $(X^{t,x}, P^{t,x}, \tilde{Z}^{t,x})$ presented above, and that $\tilde{u}$ is a decoupling field for the auxiliary FBSDE.

For $(t,m) \in [0,T] \times \mathbb{R}$ an optimal control for this example is given by the process α, defined by

$$\alpha_s := f_a^{-1}(s, M_s^{t,m}, -Z_s^{t,m}) = \tilde{Z}_s^{t,x} = \bar{a}_s V_s, \ s \in [t,T],$$

as stated in Proposition 4.1 and in Proposition 5.1 of [1], where we have used the notation aforementioned. Note that α does not depend on the initial condition (t,m). With the optimal control α we can represent the value function of this problem by

$$v(t,m) = \mathbb{E}\left[\int_t^T f\left(s, M_s^{t,m,\alpha}, \bar{a}_s V_s\right) \, \mathrm{d}s + g\left(M_T^{t,m,\alpha}\right)\right].$$

We point out that v is deterministic since we are in a Markovian setting. This especially entails the applicability of the verification theorem to show later that the antiderivative of u equals the value function v. To that end, we consider the HJB equation which takes the form

$$\begin{aligned}
0 &= -w_t(s,m) - Bmw_m(s,m) - \inf_{a\in\mathbb{R}}\left[\frac{1}{2}a^2 w_{mm}(s,m) + f(s,m,a)\right], \\
g &= w(T,\cdot),
\end{aligned} \tag{4.9}$$

(see [11, p. 46] for details). Furthermore, we define our candidate for the antiderivative of u as

$$\begin{aligned}
\widetilde{v}(t,m) &:= \int_0^m u(t,z)\,\mathrm{d}z + c(t) = \left(\frac{1}{V_t} - 1\right) m^2 + c(t) \\
&= \begin{cases} \left(l(T-t) + L\right) m^2 + c(t) & , B = 0, \\ \left(e^{2B(T-t)}\left(1 + L + \frac{l-2B}{2B}\right) - \frac{l}{2B}\right) m^2 + c(t) & , B \neq 0, \end{cases}
\end{aligned}$$

for $(t,m) \in [0,T] \times \mathbb{R}$. Our aim is to choose c such that $\tilde{v}$ solves the HJB equation.

Lemma 4.2. *If we choose*

$$c(t) := \int_t^T \overline{a}_s^2 \left(1 - V_s\right) \mathrm{d}s,$$

then the function $\tilde{v}$ is once continuously differentiable in the time variable and two times continuously differentiable in the space variable. Moreover, $\tilde{v}$ solves the HJB equation (4.9).

Proof. Since $\overline{a}$ and V are continuous, we see that $\tilde{v}(\cdot,m)$ is continuously differentiable for all $m \in \mathbb{R}$ by the fundamental theorem of calculus. In addition, we observe that $\tilde{v}(t,\cdot)$ is two times continuously differentiable for fixed $t \in [0,T]$.

Now we show that $\tilde{v}$ solves the HJB equation (4.9). Let $(t,m) \in [0,T] \times \mathbb{R}$ be fixed. We have that

$$\tilde{v}_t(t,m) = -\frac{V_t'}{V_t^2} m^2 - \bar{a}_t^2 \left(1 - V_t\right) = -\left(\frac{2B}{V_t} + l - 2B\right) m^2 - \bar{a}_t^2 \left(1 - V_t\right),$$

using that V solves the Riccati equation (4.7), and

$$\tilde{v}_m(t,m) = 2m\left(\frac{1}{V_t} - 1\right), \qquad \tilde{v}_{mm}(t,m) = 2\left(\frac{1}{V_t} - 1\right).$$

Moreover, we observe

$$\begin{aligned}
&\inf_{a\in\mathbb{R}}\left[\frac{1}{2}a^2 \tilde{v}_{mm}(t,m) + f(t,m,a)\right] \\
&\quad= \inf_{a\in\mathbb{R}}\left[a^2\left(\frac{1}{V_t} - 1\right) + lm^2 + (a - \bar{a}_t)^2\right] = \inf_{a\in\mathbb{R}}\left[\frac{1}{V_t}a^2 - 2\bar{a}_t a + \bar{a}_t^2 + lm^2\right] \\
&\quad= \frac{1}{V_t}\left(V_t \bar{a}_t\right)^2 - 2\bar{a}_t\left(V_t \bar{a}_t\right) + \bar{a}_t^2 + lm^2 = \bar{a}_t^2\left(1 - V_t\right) + lm^2.
\end{aligned}$$

Therefore, we have

$$
\begin{aligned}
&- \tilde{v}_t(t,m) - Bm\tilde{v}_m(t,m) - \inf_{a\in\mathbb{R}}\left[\frac{1}{2}a^2\tilde{v}_{mm}(t,m) + f(t,m,a)\right] \\
&\quad = \left(\frac{2B}{V_t} + l - 2B\right)m^2 + \bar{a}_t^2\left(1 - V_t\right) - 2Bm^2\left(\frac{1}{V_t} - 1\right) - \bar{a}_t^2\left(1 - V_t\right) - lm^2 = 0.
\end{aligned}
$$

Finally, the definition of $\tilde{v}$ implies $\tilde{v}(T,m) = Lm^2 = g(m)$ for all $m \in \mathbb{R}$ and thus $\tilde{v}$ solves the HJB equation (4.9). □

Lemma 4.2 immediately yields the final result.

Proposition 4.3. *We have that $\tilde{v} = v$ on $[0,T] \times \mathbb{R}$.*

Proof. Note that the function $\tilde{v}$ solves the HJB equation (4.9) according to Lemma 4.2. Additionally, for the control $\alpha_t = V_t\bar{a}_t$ there exists a solution to the SDE (4.6) according to Proposition 5.1 and Theorem 4.6 in [1]. Hence the requirements of the verification theorem (see e.g. Theorem 3.5.2 in [11]) are satisfied and the result follows. □

Proposition 4.3 eventually implies that there is an antiderivative w.r.t. m of the decoupling field u that is equal to the value function v. In particular, we observe that v is differentiable w.r.t. m and that $v_m = u$.

4.3 The general case

In the previous section we have shown in an example that the classical derivative of the value function is the decoupling field. Our goal in this section is to generalize this statement to a non-Markovian framework. We cannot, however, expect the value function v to be classically differentiable, even though we have by Proposition 4.1 the explicit representation

$$
\begin{aligned}
v(t,m) &= \mathbb{E}\left[\int_t^T f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\,\mathrm{d}s + g\left(M_T^{t,m}\right)\middle|\mathcal{F}_t\right] \\
&= \mathbb{E}\left[\int_t^T f\left(s, M_s^{t,m}, f_a^{-1}(s, M_s^{t,m}, -Z_s^{t,m})\right)\,\mathrm{d}s + g\left(M_T^{t,m}\right)\middle|\mathcal{F}_t\right], \text{ a.s.,}
\end{aligned}
$$

for all $(t,m) \in [0,T] \times \mathbb{R}$.

Consequently, we work here with weak derivatives. We show that v is weakly differentiable and that a version of the weak derivative is given by the decoupling field u. Note that we consider the version of the decoupling field provided by Proposition 4.1, and that details on weak derivatives can be found in section 2.5. Now we state the main result of this section.

Theorem 4.4. *The value function v is weakly differentiable w.r.t. m and for all $t \in [0,T]$ the weak derivative $\partial_m v$ satisfies*

$$\partial_m v(t,m) = u(t,m)$$

for $\mathbb{P} \otimes \lambda$-almost every $(\omega, m) \in \Omega \times \mathbb{R}$.

Studying the weak differentiability of v involves changes of weak differentiation, expectation and integration. Moreover, we have to apply chain rules for weak derivatives. Thus, we successively develop the weak differentiability of v using several lemmas. For these considerations we fix the time $t \in [0,T]$.

Lemma 4.5. *We have that:*

1. *For all $\omega \in \Omega$ the function $f_a^{-1}(t,\cdot,\cdot)$ is continuously differentiable and Lipschitz continuous.*
2. *The control $\hat{\alpha}$ belongs to the space $\mathcal{A}(t)$ and is weakly differentiable w.r.t. m.*
3. *It holds*
$$\operatorname*{ess\,sup}_{m\in\mathbb{R}} \mathbb{E}\left[\int_t^T (\partial_m \hat{\alpha}_s)^2 \,\mathrm{d}s\right] < \infty.$$

Proof. 1. Let $\omega \in \Omega$ be fixed. We prove that f_a^{-1} is continuously differentiable by applying the inverse function theorem. Note that the function $f(t,\cdot,\cdot)$ is two times continuously differentiable due to the assumptions (D1)-(D3). We define the continuously differentiable mapping $\varphi : \mathbb{R}^2 \to \mathbb{R}^2, (m,a) \mapsto (m, f_a(t,m,a))$, and observe that the inverse is given by the function $\psi(m,z) = (m, f_a^{-1}(t,m,z))$ for $(m,z) \in \mathbb{R}^2$, because $\psi \circ \varphi = \varphi \circ \psi = \mathrm{Id}_{\mathbb{R}^2}$. Let $(m,z) \in \mathbb{R}^2$ and $a := f_a^{-1}(t,m,z)$. The Jacobian matrix of φ at (m,a)
$$D\varphi(m,a) = \begin{pmatrix} 1 & 0 \\ f_{ma}(t,m,a) & f_{aa}(t,m,a) \end{pmatrix}$$
is invertible since $f_{aa}(t,m,a) \geq \delta_l > 0$. Hence the inverse function theorem implies that the inverse ψ is continuously differentiable in a neighbourhood of (m,z) and
$$D\psi(m,z) = \left(D\varphi(m, f_a^{-1}(t,m,z))\right)^{-1}.$$
In particular, we have
$$\frac{\partial f_a^{-1}}{\partial m}(t,m,z) = -\frac{f_{ma}(t,m,f_a^{-1}(t,m,z))}{f_{aa}(t,m,f_a^{-1}(t,m,z))} \quad \text{and} \quad \frac{\partial f_a^{-1}}{\partial a}(t,m,z) = \frac{1}{f_{aa}(t,m,f_a^{-1}(t,m,z))}.$$
The arbitrary choice of (m,z) especially yields that $f_a^{-1}(t,\cdot,\cdot)$ is continuously differentiable. The Lipschitz continuity follows since f_{ma} is bounded by the Lipschitz constant of f_a, f_m, and f_{aa} is bounded from below by δ_l (see e.g. Remark 3.1).

2. The control $\hat{\alpha}$ is given by $\hat{\alpha}_s = f_a^{-1}(s, M_s^{t,m}, -Z_s^{t,m})$ for $s \in [t,T]$. The first statement of this lemma implies that f_a^{-1} is Lipschitz continuous in the last two variables with a Lipschitz constant $L \geq 0$. Thus, we have $\hat{\alpha} \in \mathcal{A}(t)$ since
$$\begin{aligned} E\left[\int_t^T \hat{\alpha}_s^2 \,\mathrm{d}s\right] &\leq \mathbb{E}\left[\int_t^T \left(4L^2\left(|M_s^{t,m}|^2 + |Z_s^{t,m}|^2\right) + 2|f_a^{-1}(s,0,0)|^2\right) \mathrm{d}s\right] \\ &\leq \mathbb{E}\left[\int_t^T \left(4L^2\left(|M_s^{t,m}|^2 + |Z_s^{t,m}|^2\right) + 2L^2\|f_a(\cdot,\cdot,0,0)\|_\infty^2\right) \mathrm{d}s\right] < \infty. \end{aligned}$$
Here we have used that $|f_a^{-1}(s,0,0)| = |f_a^{-1}(s,0,0) - f_a^{-1}(s,0,f_a(s,0,0))|$, condition (D2) and that $M^{t,m}, Z^{t,m} \in \mathcal{H}_{t,T}^2$ by Proposition 4.1.

Moreover, the processes $M^{t,m}$ and $Z^{t,m}$ are weakly differentiable by Proposition 4.1. Therefore, we can apply the chain rule in Proposition 2.38 for almost all (ω, s) fixed and obtain that $\hat{\alpha}$ is weakly differentiable w.r.t. m with weak derivative

$$\partial_m \hat{\alpha}_s = \frac{\partial f_a^{-1}}{\partial m}(s, M_s^{t,m}, -Z_s^{t,m})\partial_m M_s^{t,m} - \frac{\partial f_a^{-1}}{\partial a}(s, M_s^{t,m}, -Z_s^{t,m})\partial_m Z_s^{t,m}, \tag{4.10}$$

for almost all $(\omega, s, m) \in \Omega \times [t,T] \times \mathbb{R}$.

3. Using equation (4.10) and the estimates on f_{ma}, f_{aa} we observe that

$$\mathbb{E}\left[\int_t^T (\partial_m \hat{\alpha}_s)^2 \,\mathrm{d}s\right] \leq \frac{2L^2}{\delta_l^2}\,\mathbb{E}\left[\int_t^T \left(\partial_m M_s^{t,m}\right)^2 \mathrm{d}s\right] + \frac{2}{\delta_l^2}\,\mathbb{E}\left[\int_t^T \left(\partial_m Z_s^{t,m}\right)^2 \mathrm{d}s\right],$$

where $L \geq 0$ is the Lipschitz constant of f_a, f_m. Finally, (4.3) yields the result. □

Lemma 4.6. *The functions $m \mapsto f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)$ and $m \mapsto g\left(M_T^{t,m}\right)$ are weakly differentiable w.r.t. m with weak derivatives*

$$\begin{aligned}\partial_m \left[f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\right] &= f_m\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\partial_m M_s^{t,m} - Z_s^{t,m}\partial_m \hat{\alpha}_s,\\ \partial_m\left[g(M_T^{t,m})\right] &= g'(M_T^{t,m})\partial_m M_T^{t,m},\end{aligned} \tag{4.11}$$

for $\mathbb{P} \otimes \lambda \otimes \lambda$-almost all $(\omega, s, m) \in \Omega \times [t,T] \times \mathbb{R}$.

Proof. This proof basically consists of the application of a chain rule for weak derivatives in the version of Corollary 2.39. First of all, we consider $f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)$. Note that $M^{t,m}$ and $\hat{\alpha}$ are weakly differentiable w.r.t. m. Moreover, $f(t, \cdot, \cdot)$ is continuously differentiable for all $(\omega, s) \in \Omega \times [t,T]$. Thus, we can apply Corollary 2.39 for almost all (ω, s) fixed and obtain

$$\begin{aligned}\partial_m\left[f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\right] &= f_m\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\partial_m M_s^{t,m} + f_a\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\partial_m \hat{\alpha}_s\\ &= f_m\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\partial_m M_s^{t,m} - Z_s^{t,m}\partial_m \hat{\alpha}_s,\end{aligned}$$

for almost every $(\omega, s, m) \in \Omega \times [t,T] \times \mathbb{R}$ by the definition of $\hat{\alpha}$.

Now we show the second equation. Note that $M_T^{t,m}$ is weakly differentiable w.r.t. m due to Proposition 4.1. Again Corollary 2.39 applied ω-wise yields

$$\partial_m\left[g(M_T^{t,m})\right] = g'(M_T^{t,m})\partial_m M_T^{t,m}, \text{ a.s.},$$

for almost every $m \in \mathbb{R}$, because g is continuously differentiable. This concludes the proof. □

Lemma 4.7. *The mapping $m \mapsto \int_t^T f\left(s, M_s^{t,m}, \hat{\alpha}_s\right) \mathrm{d}s$ is weakly differentiable w.r.t. m with weak derivative*

$$\partial_m \int_t^T f\left(s, M_s^{t,m}, \hat{\alpha}_s\right) \mathrm{d}s = \int_t^T \partial_m\left(f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\right) \mathrm{d}s, \text{ a.s.}$$

for almost every $m \in \mathbb{R}$.

Proof. Note that the function $(\omega, s, m) \mapsto f(\omega, s, M_s^{t,m}, \tilde{Z}_s^{t,m})$ is measurable as a composition of measurable functions. We aim at applying Proposition 2.36 and therefore we first check if all the requirements are met:

(1) Lemma 4.6 implies that $m \mapsto f(s, M_s^{t,m}, \hat{\alpha}_s)$ is weakly differentiable w.r.t. m.

(2) Our assumptions (D1)-(D3) ensure that the value function is well-defined. In particular, this means that for all $m \in \mathbb{R}$

$$\mathbb{E}\left[\int_t^T \left|f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\right| \,\mathrm{d}s\right] < \infty.$$

To be more precise, one can show that this expression is finite by using the mean value theorem, Lipschitz continuity of f_a, f_m and the integrability properties of $\hat{\alpha}$ and $M^{t,m}$.

(3) By Proposition 4.1 we have for all $m \in \mathbb{R}$ that $\hat{\alpha} \in \mathcal{A}(t), M^{t,m}, Z^{t,m} \in \mathcal{H}_{t,T}^2$ and

$$\operatorname*{ess\,sup}_{m\in\mathbb{R}} \mathbb{E}\left[\int_t^T \left(\partial_m M_s^{t,m}\right)^2 \,\mathrm{d}s\right] + \operatorname*{ess\,sup}_{m\in\mathbb{R}} \mathbb{E}\left[\int_t^T \left(\partial_m Z_s^{t,m}\right)^2 \,\mathrm{d}s\right] < \infty. \tag{4.12}$$

Moreover, Lemma 4.5 implies that

$$\operatorname*{ess\,sup}_{m\in\mathbb{R}} \mathbb{E}\left[\int_t^T \left(\partial_m \hat{\alpha}_s\right)^2 \,\mathrm{d}s\right] < \infty. \tag{4.13}$$

Consequently, Lemma 2.33 yields for all compact $K \subseteq \mathbb{R}$

$$\begin{aligned}\operatorname*{ess\,sup}_{m\in K} \mathbb{E}\left[\int_t^T \hat{\alpha}_s^2 \,\mathrm{d}s\right] + \operatorname*{ess\,sup}_{m\in K} \mathbb{E}\left[\int_t^T \left(M_s^{t,m}\right)^2 \,\mathrm{d}s\right]& \\ + \operatorname*{ess\,sup}_{m\in K} \mathbb{E}\left[\int_t^T \left(Z_s^{t,m}\right)^2 \,\mathrm{d}s\right] &< \infty.\end{aligned} \tag{4.14}$$

Note that by the Lipschitz continuity of f_m (see (D2)) we have for all $m \in \mathbb{R}$ that

$$f_m\left(s, M_s^{t,m}, \hat{\alpha}_s\right)^2 \leq 4L^2\left(|M_s^{t,m}|^2 + |\hat{\alpha}_s|^2\right) + 2\|f_m(\cdot,\cdot,0,0)\|_\infty^2, \text{ a.s.}, \tag{4.15}$$

for almost every $s \in [t, T]$. For any compact set $K \subseteq \mathbb{R}$ we obtain by Lemma 4.6

$$\begin{aligned}&\operatorname*{ess\,sup}_{m\in K} \mathbb{E}\left[\int_t^T \left|\partial_m\left(f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\right)\right| \,\mathrm{d}s\right] \\ &\leq \operatorname*{ess\,sup}_{m\in K} \mathbb{E}\left[\int_t^T \left(\left|f_m\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\partial_m M_s^{t,m}\right| + \left|Z_s^{t,m}\partial_m\hat{\alpha}_s\right|\right) \,\mathrm{d}s\right] \\ &\leq \operatorname*{ess\,sup}_{m\in K}\Bigg[\left(\mathbb{E}\left[\int_t^T |f_m\left(s, M_s^{t,m}, \hat{\alpha}_s\right)|^2 \,\mathrm{d}s\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\int_t^T \left(\partial_m M_s^{t,m}\right)^2 \,\mathrm{d}s\right]\right)^{\frac{1}{2}} \\ &\qquad + \left(\mathbb{E}\left[\int_t^T \left(Z_s^{t,m}\right)^2 \,\mathrm{d}s\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\int_t^T \left(\partial_m\hat{\alpha}_s\right)^2 \,\mathrm{d}s\right]\right)^{\frac{1}{2}}\Bigg] < \infty.\end{aligned} \tag{4.16}$$

The last term is finite because of the estimate in (4.15) and by (4.12)-(4.14).

Finally, we can apply Proposition 2.36 and the result follows. □

Lemma 4.8. *The value function v is weakly differentiable w.r.t. m with weak derivative*

$$\begin{aligned}\partial_m v(t,m) = \mathbb{E}\Bigg[\int_t^T \left(f_m\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\partial_m M_s^{t,m} - Z_s^{t,m}\partial_m \hat{\alpha}_s\right)\,\mathrm{d}s \\ + g'(M_T^{t,m})\partial_m M_T^{t,m}\Bigg|\mathcal{F}_t\Bigg],\end{aligned} \tag{4.17}$$

for $\mathbb{P}\otimes\lambda$-almost every $(\omega, m)\in\Omega\times\mathbb{R}$.

Proof. In order to prove that the value function v is weakly differentiable, we check if the requirements of Proposition 2.35 are met. Lemma 4.6 and Lemma 4.7 imply that the functions

$$m \mapsto \int_t^T f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\,\mathrm{d}s \qquad \text{and} \qquad m\mapsto g\left(M_T^{t,m}\right)$$

are weakly differentiable. Proposition 2.31 yields that also the sum of both functions is weakly differentiable, and again using Lemma 4.6 and Lemma 4.7, we can calculate the weak derivative explicitly

$$\begin{aligned}&\partial_m\left(\int_t^T f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\,\mathrm{d}s + g\left(M_T^{t,m}\right)\right)\\ &= \int_t^T \left(f_m\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\partial_m M_s^{t,m} - Z_s^{t,m}\partial_m\hat{\alpha}_s\right)\,\mathrm{d}s + g'(M_T^{t,m})\partial_m M_T^{t,m},\end{aligned}$$

for $\mathbb{P}\otimes\lambda$-almost every $(\omega,m)\in\Omega\times\mathbb{R}$. This implies condition (1) of Proposition 2.35. Moreover, condition (2) is fulfilled since our requirements on the functions f and g ensure that the value function is well-defined for all initial values $m\in\mathbb{R}$. To show that condition (3) is satisfied, note that by Proposition 4.1 we have $\mathbb{E}\left(M_T^{t,m}\right)^2<\infty$ for all $m\in\mathbb{R}$ and

$$\operatorname*{ess\,sup}_{m\in\mathbb{R}}\mathbb{E}\left[\left(\partial_m M_T^{t,m}\right)^2\right]<\infty. \tag{4.18}$$

Consequently, Lemma 2.33 yields that for all compact $K\subseteq\mathbb{R}$

$$\operatorname*{ess\,sup}_{m\in K}\mathbb{E}\left[\left(M_T^{t,m}\right)^2\right]<\infty. \tag{4.19}$$

Altogether, we observe for all compact $K\subseteq\mathbb{R}$

$$\begin{aligned}&\operatorname*{ess\,sup}_{m\in K}\mathbb{E}\left[\left|\partial_m\left(\int_t^T f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\,\mathrm{d}s + g\left(M_T^{t,m}\right)\right)\right|\right]\\ &\le \operatorname*{ess\,sup}_{m\in K}\left(\mathbb{E}\left[\int_t^T\left|\partial_m\left(f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\right)\right|\,\mathrm{d}s\right] + \mathbb{E}\left|g'(M_T^{t,m})\partial_m M_T^{t,m}\right|\right)\\ &\le \operatorname*{ess\,sup}_{m\in K}\left(\mathbb{E}\left[\int_t^T\left|\partial_m\left(f\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\right)\right|\,\mathrm{d}s\right] + \left(\mathbb{E}\left[g'(M_T^{t,m})^2\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\left(\partial_m M_T^{t,m}\right)^2\right]\right)^{\frac{1}{2}}\right)\\ &<\infty.\end{aligned}$$

The last expression is finite because of the estimate (cf. (D2))

$$\mathbb{E}\left[g'(M_T^{t,m})^2\right]\le 2L^2\mathbb{E}\left[(M_T^{t,m})^2\right] + 2\|g'(\cdot,0)\|_\infty^2<\infty,$$

and the equations (4.16), (4.18), (4.19). ☐

Finally, we can prove the main result of this section.

Proof of Theorem 4.4. Lemma 4.8 implies that the value function v is weakly differentiable with derivative given by equation (4.17). It holds that $Y_T^{t,m} = g'(M_T^{t,m})$ by the FBSDE (4.2) and hence we can write (4.17) as

$$\partial_m v(t,m) = \mathbb{E}\left[\int_t^T \left(f_m\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\partial_m M_s^{t,m} - Z_s^{t,m}\partial_m\hat{\alpha}_s\right)\,\mathrm{d}s + Y_T^{t,m}\partial_m M_T^{t,m}\middle|\mathcal{F}_t\right],$$

for $\mathbb{P}\otimes\lambda$-almost every $(\omega, m) \in \Omega\times\mathbb{R}$. Our goal is to expand the term $Y_T^{t,m}\partial_m M_T^{t,m}$ using the product formula. The reasoning below holds true for almost every $m \in \mathbb{R}$ fixed. Note that we can express $Y^{t,m}$ as an Itô process

$$Y_s^{t,m} = Y_t^{t,m} - \int_t^s \left(B_r Y_r^{t,m} + f_m\left(r, M_r^{t,m}, \hat{\alpha}_r\right)\right)\,\mathrm{d}r + \int_t^s Z_r^{t,m}\,\mathrm{d}W_r,\ \ s\in[t,T].$$

Moreover, the process $\partial_m M^{t,m}$ satisfies the dynamics

$$\partial_m M_s^{t,m} = 1 + \int_t^s B_r\partial_m M_r^{t,m}\,\mathrm{d}r + \int_t^s \partial_m\hat{\alpha}_r\,\mathrm{d}W_r,\ \ s\in[t,T],$$

as one can show following the same train of thought as in Lemma 3.8. Using the product formula for $Y^{t,m}\partial_m M^{t,m}$ yields

$$\begin{aligned} Y_s^{t,m}\partial_m M_s^{t,m} = Y_t^{t,m} + \int_t^s \Big[& Y_r^{t,m} B_r \partial_m M_r^{t,m} + Z_r^{t,m}\partial_m\hat{\alpha}_r \\ & - \partial_m M_r^{t,m}\left(B_r Y_r^{t,m} + f_m\left(r, M_r^{t,m}, \hat{\alpha}_r\right)\right)\Big]\,\mathrm{d}r \\ & + \int_t^T \left(Y_r^{t,m}\partial_m\hat{\alpha}_r - \partial_m M_r^{t,m} Z_r^{t,m}\right)\,\mathrm{d}W_r \\ = Y_t^{t,m} + \int_t^s & \left(Z_r^{t,m}\partial_m\hat{\alpha}_r - \partial_m M_r^{t,m} f_m\left(r, M_r^{t,m}, \hat{\alpha}_r\right)\right)\,\mathrm{d}r \\ & + \int_t^s \left(Y_r^{t,m}\partial_m\hat{\alpha}_r - \partial_m M_r^{t,m} Z_r^{t,m}\right)\,\mathrm{d}W_r,\ \ s\in[t,T]. \end{aligned} \tag{4.20}$$

Now we define an appropriate localizing sequence $(\tau_n)_{n\in\mathbb{N}}$ by

$$\tau_n := \inf\{s \geq t : |Y_s^{t,m}| \geq n \text{ or } |\partial_m M_s^{t,m}| \geq n\} \wedge T,\ \ n\in\mathbb{N}.$$

It holds $\lim_{n\to\infty}\tau_n = T$ a.s. since we have $Y^{t,m}, \partial_m M^{t,m} \in \mathcal{S}^2_{t,T}$ by Proposition 4.1. Taking the conditional expectation of both sides in (4.20) implies

$$\mathbb{E}\left[Y_{T\wedge\tau_n}^{t,m}\partial_m M_{T\wedge\tau_n}^{t,m}\middle|\mathcal{F}_t\right] = \mathbb{E}\left[Y_t^{t,m} + \int_t^{T\wedge\tau_n}\left(Z_r^{t,m}\partial_m\hat{\alpha}_r - \partial_m M_r^{t,m} f_m\left(r, M_r^{t,m}, \hat{\alpha}_r\right)\right)\,\mathrm{d}r\middle|\mathcal{F}_t\right]$$

for all $n \in \mathbb{N}$, because $\int_t^{\cdot\wedge\tau_n}(Y_r^{t,m}\partial_m\hat{\alpha}_r - \partial_m M_r^{t,m} Z_r^{t,m})\,\mathrm{d}W_r$ is a martingale on $[t,T]$. Moreover, dominated convergence implies that

$$\mathbb{E}\left[Y_T^{t,m}\partial_m M_T^{t,m}\middle|\mathcal{F}_t\right] = \mathbb{E}\left[Y_t^{t,m} + \int_t^T \left(Z_r^{t,m}\partial_m\hat{\alpha}_r - \partial_m M_r^{t,m} f_m\left(r, M_r^{t,m}, \hat{\alpha}_r\right)\right)\,\mathrm{d}r\middle|\mathcal{F}_t\right].$$

Substituting this into equation (4.17) yields

$$\begin{aligned}\partial_m v(t,m) &= \mathbb{E}\left[\int_t^T \left(f_m\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\partial_m M_s^{t,m} - Z_s^{t,m}\partial_m\hat{\alpha}_s\right)\,\mathrm{d}s\middle|\mathcal{F}_t\right] \\ &\qquad + \mathbb{E}\left[Y_t^{t,m} + \int_t^T \left(Z_s^{t,m}\partial_m\hat{\alpha}_s - \partial_m M_s^{t,m} f_m\left(s, M_s^{t,m}, \hat{\alpha}_s\right)\right)\,\mathrm{d}s\middle|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[Y_t^{t,m}|\mathcal{F}_t\right] = Y_t^{t,m} = u(t, M_t^{t,m}) = u(t,m),\ \text{a.s.},\end{aligned}$$

for almost every $m \in \mathbb{R}$. Here we have used the decoupling condition and that $Y_t^{t,m}$ is $\mathcal{F}_t$-measurable. The arbitrary choice of $t \in [0,T]$ on page 67 implies that the result is true for all $t \in [0,T]$. □

Corollary 4.9. *There exists a function $\tilde{v} : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ that satisfies:*

(1) $\tilde{v}(\omega, t, \cdot)$ is classical differentiable for almost all $\omega \in \Omega$ and all $t \in [0,T]$ and

$$\frac{\partial \tilde{v}}{\partial m}(t,m) = u(t,m),\ \textit{a.s.},$$

for all $m \in \mathbb{R}$, almost all $\omega \in \Omega$ and all $t \in [0,T]$.

(2) For all $(\omega, t) \in \Omega \times [0,T]$ we have $v(\omega,t,m) = \tilde{v}(\omega,t,m)$ for almost every $m \in \mathbb{R}$.

Proof. For all $t \in [0,T]$ and almost all $\omega \in \Omega$ fixed $v(\omega,t,\cdot)$ is weakly differentiable and thus there exists a version of v that is continuous according to Theorem 2.1.4 in [14]. We define $\tilde{v}(\omega,t,\cdot)$ to be that version. For the other $\omega \in \Omega$, that are contained in a $\mathbb{P}$-null set, we just set $\tilde{v}(\omega,t,\cdot) := v(\omega,t,\cdot)$. This construction yields that (2) is fulfilled and therefore $\tilde{v}$ is also weakly differentiable w.r.t. m with the same weak derivative, i.e.

$$\tilde{v}(t,m) = u(t,m),\ \text{a.s.},$$

for almost every $m \in \mathbb{R}$ and all $t \in [0,T]$. Now applying Theorem 2.42 for all $t \in [0,T]$ and almost all $\omega \in \Omega$ fixed implies (1), since $\tilde{v}(\omega,t,\cdot)$ and $u(\omega,t,\cdot)$ are continuous. □

Remark 4.10. Corollary 4.9 implies that the value function has a version that is continuously differentiable in m with the decoupling field being its derivative. If the value function is already continuous, the above statement would be true for $\tilde{v} = v$. We suspect that this is indeed true, but further research is required. In particular, one has to show that the value function is continuous w.r.t. m in our non-Markovian setting.

References

[1] S. Ankirchner and A. Fromm. “Optimal Control of Diffusion Coefficients via Decoupling Fields”. In: *SIAM Journal on Control and Optimization* 56.4 (2018), pp. 2959–2976.

[2] R. Cairoli and R.C. Dalang. *Sequential stochastic optimization*. Vol. 238. John Wiley & Sons, 2011.

[3] I. Ekren and J. Zhang. “Pseudo-Markovian viscosity solutions of fully nonlinear degenerate PPDEs”. In: *Probability, Uncertainty and Quantitative Risk* 1.1 (2016).

[4] I. Ekren, N. Touzi, J. Zhang, et al. “Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part II”. In: *The Annals of Probability* 44.4 (2016), pp. 2507–2553.

[5] N. El Karoui, S. Peng, and M. C. Quenez. “Backward stochastic differential equations in finance”. In: *Mathematical finance* 7.1 (1997), pp. 1–71.

[6] A. Fromm. “Theory and applications of decoupling fields for forward-backward stochastic differential equations”. PhD thesis. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, 2015.

[7] N. Kazamaki. *Continuous exponential martingales and BMO*. Springer, 2006.

[8] M. Kobylanski et al. “Backward stochastic differential equations and partial differential equations with quadratic growth”. In: *the Annals of Probability* 28.2 (2000), pp. 558–602.

[9] J. Ma, J-M Morel, and J. Yong. *Forward-backward stochastic differential equations and their applications*. 1702. Springer Science & Business Media, 1999.

[10] V. G. Maz’ia. *Sobolev Spaces: With Applications to Elliptic Partial Differential Equations*. Springer, 2011.

[11] H. Pham. *Continuous-time stochastic control and optimization with financial applications*. Vol. 61. Springer Science & Business Media, 2009.

[12] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*. Vol. 293. Springer Science & Business Media, 2013.

[13] J. Yong and X. Y. Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*. Vol. 43. Springer Science & Business Media, 1999.

[14] W. P. Ziemer. *Weakly differentiable functions: Sobolev spaces and functions of bounded variation*. Vol. 120. Springer Science & Business Media, 1989.

Selbstständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.
Seitens des Verfassers bestehen keine Einwände die vorliegende Masterarbeit für die öffentliche Benutzung im Universitätsarchiv zur Verfügung zu stellen.

Jena, den 4. April 2019 Julian Wendt